THE UNIVERSITY *of* EDINBURGH

Comparative Proteomic Analysis of *Clostridium difficile*

by

Caroline Hazel Chilton

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy by Thesis

The University of Edinburgh

College of Medicine and Veterinary Medicine

December 2010

The candidate confirms that this thesis has been composed by herself, that the work submitted is her own, and that appropriate credit has been given where reference has been made to the work of others. This work has not been submitted for any other degree or professional qualification.

Signed .....

Date .....

Abstract

The recent increase in availability of next generation sequencing methodologies has led to extensive analysis of the genome of *Clostridium difficile*. In contrast, protein expression analysis, crucial to the elucidation of mechanisms of disease, has severely lagged behind. In this study, in-depth proteomic analysis of three strains of varying virulence, demonstrated previously in an animal model, has been undertaken against a background of the sequenced genomes. Strain B-1 is a historic, virulent, ribotype 005 clone, strain A represents the emerging hypervirulent 027 ribotype, while strain Tra5/5, ribotype 001, is of low virulence.

To undertake a comprehensive overview of the expressed proteome, both 1D and 2D gel electrophoresis were used to separate and display the protein content of each isolate. This was coupled to MALDI-TOF and LC-MS/MS mass spectrometry for protein identification. A total of 888 different proteins were characterised by comparative analysis of isolates grown in parallel for 64 hours on blood agar. Of these, only 38% were shared between all isolates. An additional 350, 243 and 398 proteins were detected from broth cultures, and the use of a hexapeptide bead library, designed to capture low abundance proteins, led to the detection of a further 148, 127, and 171 proteins in strains A, B-1 and Tra5/5 respectively.

Relative differential expression was investigated using Differential In Gel Electrophoresis (DIGE), and five proteins were shown to have a statistically higher concentration in strain A, twelve in strain B-1 and eight in strain Tra5/5. A number of these were surface proteins, with selected S-layer proteins found to

be up-regulated in each strain, and the flagellar protein, FliC, up-regulated in both A and B-1. Furthermore, differential post-translation modification events were seen in flagellar and S-layer proteins.

In-vivo expression of these proteins was mapped using Western blotting. Immunodetection of the majority of these, including FliC and the high molecular weight S-layer protein, were conserved between the three strains, but a notable series of immunoreactive protein spots were present in strains A and Tra5/5 but not B-1, most likely corresponding to an additional S-layer protein present in the genomes these strains, but not that of B-1.

Protein expression differences for a number of previously proposed virulence proteins were evident between strains, including toxin B, sporulation, flagella and the S-layer proteins, metabolic enzymes, stress response proteins and ABC transporters. This study strongly supports the view that the virulence of *Clostridium difficile* is multifactorial, and that a number of related factors, although not directly required for pathogenicity, may serve to modulate the virulence of individual strains.

Acknowledgements

This project was funded by the Health Protection Agency.

First and foremost, I would like to thank my supervisors for all their help, support and encouragement. Professor Haroun Shah for his constant enthusiasm and for interesting me in *Clostridium difficile* in the first place, Professor Saheer Gharbia for all her ideas and guidance and for scaring me into finishing the thesis in six weeks, Professor Ian Poxton for his invaluable advice and for reminding me not to carried away by technology, and Professor Pete Borriello for his expertise throughout the project and thorough dissection of the (almost) finished manuscript.

I must also extend my gratitude to Dr Luc Guerrier for his input and assistance with the ProteoMiner work, and to Professor Neil Fairweather for taking so much time to advise and help me with western blotting issues.

Sincere thanks must go to all my colleagues and friends in DBHT who have made the past three years so enjoyable, as well as offering a huge and on-going range of help, advice and cake.

I thank all my fantastic friends, especially my sisters Linda and Rosie, for listening to my gel-related rants, musings, problems, despairs and triumphs throughout the project. It was probably excruciating for you all, but was invaluable to me!

Finally I would like to thank the rest of my family. Without the foresight and loving support of my grandparents, I would never have got this far, and I will always be grateful. Above all, I would like to thanks my parents who have always been there with love, understanding, and cups of tea in all situations. I would particularly like to thank my father, John, for his unexpected and detailed proof reading of my manuscript, and I dedicate this thesis to my mother, Dr Janet Haresnape, who will always be my inspiration.

Contents

Figures

Tables

Chapter 1 – Introduction

1.1 – *Clostridium difficile*

Clostridium difficile is a nosocomial pathogen responsible for a range of gastrointestinal associated symptoms including severe diarrhoea and pseudomembranous colitis (PMC) (Larson *et al*, 1978; Kelly and Lamont, 1998). It was first identified in 1935 from the stool samples of healthy new born infants (Hall and O'Toole, 1935). However, it was not until nearly 40 years later that it was confirmed as pathogenic, when two groups simultaneously identified it as the aetiological agent of antibiotic-associated pseudomembranous colitis (PMC)(Bartlett *et al*, 1978; Larson *et al*, 1978). In the late 1980s, *C. difficile* emerged as a problematic hospital acquired infectious agent, frequently associated with antibiotic treatment (Larson and Borriello, 1990). In recent years, there has been an increase in number and severity of nosocomial infections, as well as an increase in community care acquired infections, leading to this pathogen becoming an important focus for research (Rupnik *et al*, 2009).

1.1.1 – Pathogenicity

C. difficile is a rod-shaped, Gram-positive, anaerobic, spore-forming bacterium. Although several putative virulence factors have been identified, two major toxins, identified in the early 80s (Bartlett *et al*, 1980) and produced by many *C. difficile* strains are thought to be the main cause of infection associated pseudomembranous colitis. Toxin A (308kDa) and toxin B (270kDa) show high sequence and functional homology to each other, and to other large clostridial

toxins, and are among the largest bacterial toxins known (Voth and Ballard, 2005). Both toxin A and toxin B have been shown to be cytotoxic, with toxin B showing a 1000-fold greater cytotoxicity than toxin A (Sullivan *et al*, 1982), although their respective roles in the disease causing mechanism have long been disputed.

Originally, toxin A was thought to be the main cause of pathogenicity, as it was shown to cause haemorrhagic fluid secretion in ligated rabbit ileal loops where toxin B did not (Mitchell *et al*, 1986; Lima *et al*, 1988).An early study by Lyerly *et al* suggested that although toxin B alone was insufficient to cause disease in healthy intestinal tissue, it may gain access to underlying tissues in previously damaged areas, and exacerbate disease (Lyerly *et al*, 1985). Subsequently however, toxin A-negative, toxin B-positive strains that are capable of causing fluid accumulation in ligated ilea l loop assays and disease in the hamster model have been identified (Borriello *et al*, 1992), and have been isolated from patients with *C. difficile* infection (CDI) (Alfa *et al*, 2000). The role of toxin B has remained controversial, with both toxins being shown to independently disrupt T84 human intestinal cells (Hecht *et al*, 1992). More recently, work by Lyras *et al* (2009), both in a hamster model and in-vitro cell lines, showed that it is toxin B, not toxin A, that is essential for virulence (Lyras *et al*, 2009). However, Kuehne *et al* (2010) showed that strains producing both toxins together, or either toxin alone, were virulent. Hence the mechanisms of interaction of toxin A and toxin B still remain unclear.

The two toxins have similar modes of action, mediating their effects through glucosylation of small GTP binding proteins such as Rho, Rac and Cdc42, thereby interrupting actin cytoskeleton assembly (Borriello, 1998, Voth and Ballard, 2005). This leads to loss of cell structural integrity as well as interrupting GTPase signalling pathways and triggering apoptosis.

The two toxin genes, *tcdA* and *tcdB* are located on a 19.6kb genetic locus know as the pathogenicity locus (PaLoc) (Figure 1.1). This region also encodes three additional open reading frames, *tcdR, tcdE* and *tcdC* and has been found to be highly conserved between toxigenic strains (Hammond and Johnson, 1995).

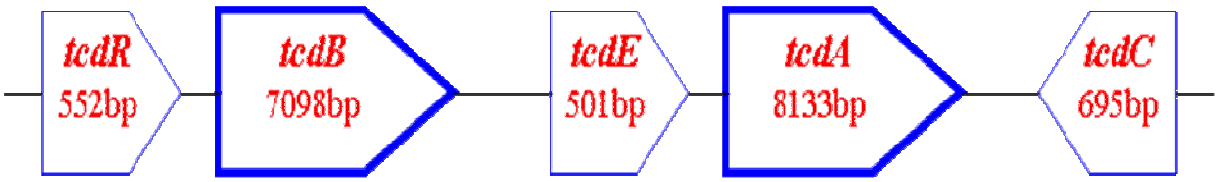


Figure 1.1 – *The Pathogenicity locus (PathLoc) of Clostridium difficile (adapted from Poxton et al, 2001). TcdB and tcdA encode toxins B and A respectively. TcdR encodes a positive regulator of toxin production, tcdC encodes a negative regulator of toxin production, and tcdE encodes a holin like protein.*

Hundsberger *et al* established that transcription of *tcdC* is high during early exponential growth phase, whereas transcription of *tcdR, tcdB, tcd E and tcdA* is low at this point, and high during late stationary phase leading to the suggestion that *tcdC* encodes a negative regulator of the toxin genes (Hundsberger *et al*, 1997). *TcdR* is a small open reading frame upstream of the two toxin genes encoding a 22kDa protein that shows sequence identity to DNA

binding proteins (Moncrief *et al*, 1997). Moncrief *et al* investigated the promoter functionality of this protein, and suggested it acts as a positive regulator of toxin production. More recently, *tcdE* has been shown to exhibit homology to holins (Tan *et al*, 2001) and may act as a lytic protein to facilitate the release of the toxins from *C. difficile*.

It has been shown that emerging 027 clone strains exhibit deletions in the *tcdC* gene leading to truncation of the TcdC protein, which may cause the observed increase in toxin production in epidemic 027 strains (Deneve *et al*, 2009; Warny *et al*, 2005). However, the mechanism of toxin regulation is not fully understood, and many other factors appear to affect toxin production (Dupuy *et al*, 2008). In addition, the large variation in the toxinotypes of pathogenic *C. difficile* strains indicates that toxin production is only one aspect of virulence, and that many interacting pathways play a role in the pathogenicity of this organism, and so the increased severity of the emerging strains is likely to be multifactorial (Rupnik, 2008).

Some strains of *C. difficile* produce an additional binary toxin with actin-specific ADP-ribosyltransferase activity. First identified in 1988 (Popoff *et al*, 1988), the prevalence of binary toxin in disease causing *C. difficile* isolates varies (Geric *et al*, 2003; Spigaglia and Mastrantonio, 2004), but it has been suggested that CDAD caused by binary toxin producing strains is more frequently associated with community-acquired infection and increased abdominal pain (Barbut *et al*, 2005). A recent report by Schwan *et al* (2009) (Schwan *et al*, 2009)has suggested that binary toxin induces microtubule proliferation. These form protrusions at

the epithelial cell surface which interact with the bacterial cells, thereby increasing adherence. Epidemic 027 ribotype strains have been shown to harbour the binary toxin encoding genes (Stabler *et al*, 2009) and produce detectable levels of the toxin (Warny *et al*, 2005).

The identity and roles of the other *C. difficile* virulence factors are less well understood. *Clostridium difficile* has the ability to produce endospores, which play a pivotal role in the transmission of disease, and so sporulation, and subsequent germination processes, may also be key for virulence. Bacterial spores are resistant to many chemical and physical agents including heat, high pressure, desiccation, UV, acids and alcohols (Setlow, 2007). This allows them to persist within a hospital environment, and making outbreaks increasingly difficult to control. In addition, *C. difficile* spores have been shown to be resistant to many hospital cleaning products, some of which appear to increase sporulation of epidemic strains (Fawley *et al*, 2007). It has been suggested that hypervirulent 027 ribotypes produce significantly more spores, earlier, than non 027 strains (Merrigan *et al*, 2010) and to have an increased sporulation rate (Akerlund *et al*, 2008). Sporulation and germination mechanisms for the model organism *Bacillus subtilis* are well established, but less well understood in Clostridial species. The sporulation initiation pathway has recently been investigated (Underwood *et al*, 2009), and differences between the pathways of *B. subtilis* and *C. difficile* shown, confirming that although bacillus is a useful model, it cannot be used to directly draw conclusions about the importance of sporulation and germination factors in *C. difficile*.

Adhesion is a critical early step in host colonisation, and it has been demonstrated that different *C. difficile* strains of varying virulence exhibit variation in their ability to colonise and adhere to the gastrointestinal tract in the hamster model (Borriello *et al*, 1988b). This has been disputed however, and Waligora *et al* found no difference in the adherence of toxigenic and non-toxigenic strains, suggesting that earlier observed variation was in mucosal association rather than true cell adherence (Waligora *et al*, 1999). .A number of cell surface proteins coordinating adhesion of the bacterial cell to the gut wall have been identified, although the full mechanism has not been elucidated. These include adhesins such as Cwp66 (Waligora *et al*, 2001), surface layer (S-layer) proteins, cell wall proteins and the flagella as well as a number of S-layer protein paralogs (Wright *et al*, 2005). Stress response proteins (Henderson *et al*, 2006) and hydrolytic and proteolytic enzymes (Seddon *et al*, 1990) have also been linked to virulence.

1.1.2 – Epidemiology

In recent years, the rate and severity of CDI has been increasing due to a previously uncommon 027 ribotype. Such strains are also classified as restriction endonuclease analysis type BI, and North American pulsed field gel electrophoresis type NAP1, and have caused geographically dispersed outbreaks of severe CDAD globally (McDonald *et al*, 2005) . Ribotype 027 was first linked with severe outbreaks in Quebec in the early 2000s (Loo, 2006) but was uncommon before 2000 (Hubert *et al*, 2007), accounting for fewer than 1% of US *C. difficile* isolates (McDonald *et al*, 2005). This ribotype has now been detected in all Canadian provinces, and at least 40 states in the US (O'Connor *et*

al, 2009). In addition, by 2008, this 027 ribotype had been detected in 16 European countries, causing major outbreaks in many including Belgium, Germany, Finland, France, Ireland, Luxembourg, the Netherlands, Switzerland and the United Kingdom (Kuijper *et al*, 2008; Freeman *et al*, 2010)

The first major outbreak in the UK caused by ribotype 027 was at Stoke-Mandeville hospital between October 2003 and June 2004, and involved 174 cases (Anonymous, 2005; Smith A., 2005). A second outbreak occurred between October 2004 and June 2005 involving 160 new cases. The large increase in CDI in recent years has put a huge strain on hospitals and residential care facilities. The large outbreaks at Stoke-Mandeville hospital led to an investigation by the United Kingdom Healthcare Commission which made a number of recommendations, including changes in the reporting of *C. difficile* cases. In England, reporting of *Clostridium difficile* in patients over 65 became mandatory in 2004, and this was extended to include all patients over the age of two in April 2005. This enhanced surveillance has helped to combat the rise in cases, with a marked decrease in the number of cases reported since 2007 (Freeman *et al*, 2010).

1.1.3 – Typing of *C. difficile*

Various typing methods have been described to distinguish between different *C. difficile* strains, however, no universally accepted method for typing, and subsequent nomenclature, of *C. difficile* isolates exists. The vast array of different typing methods available complicates the picture and there is therefore

a need for a common typing method which is reproducible across different laboratories.

Originally, typing methods focused on phenotypic properties of different strains, most notably serogrouping (Delmee *et al* , 1985). Other phenotypic methods include phage typing (Sell *et al*, 1983) and whole cell pyrolysis mass spectrometry analysis (Cartmill *et al*, 1992). Immunochemical fingerprinting of the EDTA extracted surface proteins was successfully used by Poxton *et al* (1984) to confirm *C. difficile* as an infectious agent, and that the spread of a single strain was responsible for an outbreak. However, as molecular methodologies have improved, molecular typing systems have been preferentially used.

The gold standard for bacterial typing is generally pulsed field gel electrophoresis, usually in conjunction with restriction endonucleases (Li *et al*, 2009). Chromosomal DNA digests are prepared with an infrequently cutting restriction enzyme, to give 10-20 fragments. These fragments differ in size between strains, and gel electrophoresis separates the resulting fragments for strain identification. Pulsed field GE applies alternating electric fields at different angles, increasing the resolution of bands to give sufficient sensitivity to distinguish between small size differentials of large DNA fragments making this a highly discriminatory typing technique. However, problems have been reported in using this technique for *C. difficile*, particularly due to degradation of DNA for several strains from serogroup G (Kristjansson *et al,* 1994; Bidet *et al*, 2000).

Restriction enzymes are also used for typing by restriction endonuclease analysis (REA). In this technique, frequently cutting restriction enzymes are used, generating hundreds of fragments, which are then resolved by agarose gel electrophoresis. This has been shown to be highly discriminatory and universally applicable (Clabots *et al*, 1993), but the resulting gel pattern is highly complex and strain identification by visual assessment is difficult.

Another commonly used method is PCR ribotyping, which distinguishes strains according to polymorphisms in the 16S-32S rRNA spacer region. Gurtler (1993) first used PCR to amplify this spacer region in *C. difficile*, and produced DNA fingerprints which distinguished 14 different ribotypes from a set of 24 isolates. This method was subsequently modified by O'Neill *et al* (1996) who simplified the extraction process and modified the primers used for PCR, to allow this method to be used for routine typing of *C. difficile* isolates. It was suggested that due to the speed and simplicity of this method, and the problems with DNA degradation during PFGE, ribotyping is the most discriminatory and reproducible method for typing *C. difficile* isolates (Bidet *et al*, 2000). However, PCR –ribotyping still did not discriminate between isolates within some serogroups, and reproducibility of the agarose gel electrophoresis between laboratories has been low. However, in 2008, capillary gel sequencing was used to develop this technique further (Indra *et al*, 2008). This use of capillary gel sequencing instead of agarose gel electrophoresis, removes many of the problems associated with comparison of typing results between different laboratories, increasing the potential of ribotyping as a universal tool for typing *C. difficile* isolates.

A centrally funded *Clostridium difficile* Ribotyping Network is available for England and Northern Ireland, allowing rapid typing of strains responsible for *Clostridium difficile* associated disease (CDAD) cases (http://www.hpa.org.uk/ProductsServices/InfectiousDiseases/LaboratoriesAndReferenceFacilities/ClostridiumDifficileRibotypingNetworkForEngland/). This has enabled monitoring of strains responsible for epidemic outbreaks, endemic infection, and reoccurrence of disease, and allowed the changing epidemiology of *C. difficile* infections to be scrutinised. The emergence of highly infectious, previously uncommon, 027 ribotype strains have been linked to the epidemic spread of infection (McDonald *et al*, 2005).

Multiple-locus variable number tandem repeat analysis (MLVA) is a technique that uses PCR amplification of specific loci, which contain repeated DNA sequences that vary in copy number between strains (Van den Berg *et al*, 2007). These regions rapidly evolve, so this technique can be particularly useful in mapping outbreaks.

A number of other techniques have also been used for typing, including toxinotyping (Rupnik *et al*, 1998), where strains are distinguished based on polymorphisms in the *tcdA* and *tcdB* toxin genes, and amplified fragment length polymorphism (AFLP) (Klaassen *et al*, 2002), where whole cell DNA is digested with specific restriction endonucleases, and restriction site-specific PCR amplification used to generate profiles distinguishable between strains.

The vast array of typing techniques complicates the epidemiological picture of *C. difficile* infection. Different techniques have different pros and cons, and are often useful in different situations. For example, rapid methods such as MLVA or AFLP are good for monitoring outbreak situations, where as more reproducible methods such as PCR ribotyping or PFGE are required for larger scale or multi-centre epidemiological studies. As more worldwide studies into this complicated pathogen are being undertaken, the need for a universally reproducible typing method is growing.

The genome of a *Clostridium difficile* reference strain, known as 630, was sequenced in 2006 (Sebaihia *et al*, 2006). Since then, in order to better understand reasons for the variation in virulence of strains, the genomes of other strains have also been sequenced including a virulent 027 ribotype (Stabler, 2009) and the three isolates of varying virulence selected for genome sequencing at the HPA and used in this project (unpublished data). However, this has given only limited insight into factors responsible for varying virulence. Many of the differences between strains with differential pathogenic profiles are subtle, and are likely to be manifested in the expression of potential virulence factors. Microarray analysis has also been used to identify elements of the genome being actively transcribed (Sebaihia *et al*, 2006;) which gives an additional level of information as to the genes being expressed within a given strain. However, due to mRNA turnover and degradation, this does not necessarily correspond to levels of proteins present, and cannot be used to quantify differences in levels of protein expression, which can only be done by investigating the proteome. Therefore, the use of proteomic techniques to

investigate protein expression differences is of particular value in identifying factors which may be responsible for variation on virulence profiles of different isolates. In order to fully understand the variation between strains, and elucidate factors responsible for increased virulence of some strains, the use of a combination of techniques is invaluable. This study aims to increase understanding of proteomic variation between different *C. difficile* strains of varying virulence.

1.2 - Proteomics

Proteomics can be defined as the large-scale investigation of all the proteins within a particular sample, particularly their structure and function. The term 'proteomics' was coined in 1997 as an analogy of 'genomics' for the large-scale study of genes (James, 1997) where the 'proteome' is the entire complement of proteins in a given sample (Wilkins *et al*, 1996) . The proteome, unlike the genome, is a highly dynamic concept, varying enormously according to tissue, stage of growth, environmental conditions, and disease state, so by definition, an organism will have many different 'proteomes' throughout its lifecycle.

One of the great advantages of proteomics is that it gives an insight into the real-time biochemical changes occurring within a cell, which cannot be determined from genetic data. For example, it can be used to monitor changes within a cell over time, where different proteins are expressed at different points on a growth curve, but the genetic information remains constant. Proteomics can therefore be used in microbiology to monitor subtle expression differences between different strains of the same organism which have very similar genetic makeup.

In recent years, proteomics has been increasingly used in the field of microbiology, particularly in microbial characterisation (Shah, 2010). Pathogenic strains which express virulence factors can, for example, be distinguished from non-pathogenic strains which do not. Proteomics can also be used to identify different virulence factors in strains which have differing pathogenic potential, but very similar genome sequences.

Mass Spectrometry (MS) and the subsequent comparison of peptide mass fingerprints against a comprehensive database have become firmly established as successful proteomics techniques for the identification of unknown proteins. However, in order to be used successfully for complex samples such as microbial cell protein extracts, the upstream protein separation techniques are crucial. Common methods of separation prior to MS identification include chromatography and polyacrylamide gel electrophoresis (PAGE).

1.3 - Proteomic techniques

1.3.1 Protein identification

PAGE based methods have been well described as tools for bacterial proteomics (Antelmann and Hecker, 2010). Two dimensional gel electrophoresis allows separation and visualisation of large numbers of proteins from a complex protein sample such as a bacterial whole cell protein extract. Proteins are separated in the first dimension according to isoelectric point (pI) and in the second dimension according to mass, and visualised by staining. Individual

proteins are seen as separate spots on the gel, and can be excised for trypsin digestion and protein identification by mass spectrometry, usually Matrix Assisted Laser Desorption Ionisation – Time of Flight (MALDI-TOF) MS. This gives rise to a so called 'reference map' with large numbers of proteins mapped and identified, and allows comparison of different samples. Two dimensional reference mapping allows identification of only 10 – 30% of the most abundant proteins in the proteome (Cordwell *et al*, 2001), but has been successfully employed for proteomic analysis of a large range of bacteria (Encheva *et al*, 2005), (Bernardini *et al*, 2007a); (Riedel *et al*, 2006); (Eymann *et al*, 2004), including the *Clostridium* species *C. perfringens* (Alam *et al*, 2009) and *C. acetobutylicum* (Sullivan and Bennett, 2006).

More sensitive mass spectrometry methodologies, such as Liquid Chromatography quadrapole mass spectrometry (LC-MS/MS), can be used to identify a large number of different proteins from a complex peptide mixture. Separation prior to digestion is then less critical, so LC-MS/MS can be coupled with 1D gel electrophoresis, instead of requiring the time consuming and technically difficult 2D gel electrophoresis. The increased sensitivity of the mass spectrometry methods allows proteins expressed at levels below that visualised on 2D gels to be indentified (Cordwell *et al*, 2001). Moreover, it does not limit the pI range of the proteins which can be identified as is the case with 2D gel electrophoresis.

1.3.2 Differential in gel electrophoresis

Two dimensional reference mapping can be used to separate and identify many of the components of complex protein samples. However, direct comparison of 2D gel images is problematic. Variation in electric and pH fields, temperatures and gel polymerisation can affect protein migration, which makes meaningful comparisons of 2D profiles of different samples difficult. In order to overcome this problem, different protein samples can be labelled with different fluorescent dyes, and run together on the same 2D gel. This eliminates gel to gel differences and allows direct comparison and relative quantification of samples.

Fluorescent labelling of protein samples prior to 2D electrophoresis was already an established technique, but the development of the cyanine dyes used in DIGE analysis allowed differently labelled samples to be run on the same gel (Unlu *et al*, 1997). These CyDyes are matched for size and pI and do not disturb the relative migration of proteins during electrophoresis, but are fluorescently distinct, allowing separate visualisation of each labelled sample. They are also more sensitive than many post-staining techniques, so allow visualisation of lower abundance proteins (Karp *et al*, 2008). This technology has been successfully used in a number of areas, including comparisons of bacterial pathogens such as salmonella (Encheva *et al*, 2009), and listeria (Van de Velde *et al*, 2009).

1.3.3 ProteoMiner treatment

One of the major drawbacks of the classical proteomic methodology involving 2D gel electrophoresis followed by identification using mass spectrometry is the relatively small proportion of the proteome it reveals (Cordwell *et al*, 2001)

(Righetti *et al*, 2003). Only the more abundant proteins can be visualised on 2D gels, and identification by MALDI-TOF mass spectrometry requires a relatively large amount of protein, therefore much of the so-called 'deep proteome', comprising the less abundant proteins, remains obscure. The use of more sensitive LC-MS/MS techniques increases coverage, but many proteins still go undetected. Many potential virulence factors are thought to be expressed at low levels and hence masked by the relatively large levels of a few, highly abundant, protein components. Therefore any insights into variations in the deep proteome, which can comprise up to 50% of the total proteome (Righetti and Boschetti, 2008), would be of particular interest.

A variety of pre-fractionation techniques can be used to increase the coverage of the proteome, or to target specific sub-sets of proteins. These include the classic centrifugal cell-fractionation scheme as well as a range of chromatographic and electrophoretic separation techniques, targeting techniques, and abundant protein depletion tools.

One of the most commonly used depletion tools is the immuno-subtraction of most abundant species via antibody columns (Pieper *et al*, 2003). These have been most successfully used for serum or plasma samples where a small number of highly abundant protein species constitute the majority of the protein present, masking many of the proteins present at lower abundance (Echan *et al*, 2005; Pernemalm *et al*, 2009), but are less useful for complex bacterial protein extracts where commercial columns are not available. Other affinity capture methods target specific subsets of lower abundance proteins.

Examples include the capture of glyco-proteins using lectins (Hirabayashi, 2008; Pernemalm *et al*, 2009), capture of phosphoproteins using iron affinity (Sun *et al*, 2005), and pair-wise protein-protein interactions, such as in yeast 2-hybrid systems (Li *et al*, 2004).

Chromatographic approaches extensively explored by Fountoulakis' group include affinity chromatography on heparin gels (Fountoulakis and Takacs, 1998), hydrophobic interaction chromatography (Fountoulakis *et al*, 1999b) and hydroxyapatite chromatography (Fountoulakis *et al*, 1999a). Electrophoretic pre-fractionation techniques include acrylamide gel based methods such as the classical one and two dimensional gel electrophoresis, and gel-free isoelectric focusing.

Many of these methods have major drawbacks. Immunodepletion runs the risk of co-depletion of minor proteins alongside the affinity ligands. Column chromatography relies on the availability of the relevant columns, and elution of samples from chromatographic columns requires huge amounts of salts, so eluted fractions are often dilute and require concentration steps. Moreover, they are incompatible with downstream IEF/IPG first dimensional steps due to the high salt concentrations. Pre-fractionation allows detection of a greater proportion of proteins, but within each fraction, there still exists the inherent problem that low abundance proteins are masked by those present at higher concentrations, so complex protein samples of high dynamic range cannot be fully explored.

A different approach to solving this problem utilises a combinatorial library of hexapeptide ligands to capture and 'amplify' the low abundant proteins while simultaneously reducing the relative concentration of abundant species (Boschetti and Righetti, 2008). The concept of immobilising a peptide library on beads in a 'one bead, one peptide' approach dates back to 1991 when Lam *et al* (Lam *et al*, 1991) described the development of such a library to investigate ligand-acceptor binding affinities. This approach was only more recently applied to the field of proteomics and the exploration of the low abundant proteome (Thulasiraman *et al*, 2005) and the supporting technology is now being developed by BioRad and is commercially available under the name 'ProteoMiner'.

The ProteoMiner technology comprises a library of hexapeptide ligands immobilised on resin beads. These are synthesised using a 'split, couple, recombine' solid-phase peptide synthesis methodology pioneered by Merrifield (Merrifield, 1965), in which resin beads are split into fractions, and each fraction of beads coupled to a different amino acid before the fractions are pooled and randomised. These steps are repeated to give the required number of amino acids per peptide. Using the 20 natural amino acids, and 6 coupling steps, this will create a hexapeptide library containing 20^6, or 64 million, different combinations of peptide ligand (Thulasiraman *et al*, 2005). Due to the hugely diverse nature of protein-protein interactions, such a vast library would be expected to contain a ligand which binds specifically to each protein in any given sample. The beads containing ligands binding to the most abundant proteins will rapidly become saturated, and excess protein will flow through

and be discarded, whereas trace proteins will continue to bind to the corresponding immobilised ligand, which may never reach saturation. The bound proteins are eluted from the beads, resulting in a protein sample in which the concentration of the low abundant proteins has been greatly increased in comparison to the highly abundant proteins. This technique has been successfully used to identify a number of additional proteins in biological samples such as urine (Castagna *et al*, 2005), serum (Sennels *et al*, 2007), and bile (Guerrier *et al*, 2007) as well as to identify trace impurities on natural and recombinant biopharmaceuticals (Fortis *et al*, 2006).

1.3.4 Western blotting

Clostridium difficile is a highly ubiquitous organism, and its spores are found in a wide range of habitats within the environment. However, it is the interaction of the organism with the host that mediates its virulence and pathogenicity. A number of the potential virulence factors identified in the literature are proteins mediating bacterial-host interactions such as adhesion (Pechine *et al*, 2005a). Hence, some of the most interesting subsets of the *C. difficile* proteome are those proteins interacting with the host cells, which are often targets for the host immune systems.

Western blotting using host serum can identify those proteins from a complex bacterial extract which are recognised by the host serum antibodies, and so identify proteins potentially involved in interaction with the host cells. Two dimensional Western blotting has been successfully used previously to identify

immunoreactive cell surface proteins in *Clostridium difficile* (Wright *et al*, 2008), but has not been used previously with whole cell protein extracts.

1.4 - Aims and Objectives of the present study

While there has been a wealth of information on the genome of *Clostridium difficile*, the proteome by contrast has not been widely characterised. Recent studies have looked at the cell surface proteins of the 630 reference strain (Wright *et al*, 2005), and analysed the proteins from strain 630 spores (Lawley *et al*, 2009). However, no analysis of cellular proteins, or comparative analysis of proteomic differences between isolates of varying virulence has so far been described, and was therefore a major aim of this study.

Three strains were selected for comparative analysis, belonging to different ribotypes, and with different pathogenic potentials. Strain A is a highly virulent clone isolated from the 2005 Stoke-Mandeville outbreak, and was chosen as an emerging epidemic 027 ribotype. Strain B-1 is a highly virulent, ribotype 005 clone isolated from a patient with pseudomembranous colitis in the early 1980s, whereas strain Tra5/5, a 001 ribotype clone, is much less virulent in the hamster model, and was isolated from the faeces of a healthy infant. The 630 reference strain was also used for comparison in some instances, as it has been well characterised in the literature. These strains were selected to reflect evolutionary changes of this pathogen over time, as well as allowing investigation of differences between isolates of varying virulence.

The genomes of the three strains chosen for this analysis have recently been sequenced within the Department for Bioanalysis and Horizon Technologies at the Health Protection Agency Centre for Infections (unpublished data). Therefore another aim of the present study was to map the proteomic data to the genetic information to gain a better insight into the biology of this species.

To undertake this work, two separate approaches were used, firstly 1D SDS-PAGE coupled with LC-MS/MS using an LTQ-orbitrap (Thermo Fisher Scientific, San Jose, USA), and secondly 2D SDS-PAGE coupled with MALDI-TOF MS using a MALDI-TOF *Reflectron* (Waters, Milford, USA). Differential in gel electrophoresis (DIGE) was used to compare the reference maps, and give some relative quantification of expression, and the ProteoMiner beads were used to attempt to increase coverage of the low abundance proteome. In addition, Western blotting of the protein extracts with serum samples was carried out to identify proteins that illicit an antibody immune response.

Chapter 2 – Materials and Methods

2.1 – Optimisation of Protein Extraction

2.1.1 – Bacterial Strains

Three *Clostridium difficile* isolates with varying biological and virulence spectra were used in this study alongside the 630 reference strain (Sebaihia *et al*, 2006).

Strain B-1 was isolated from a patient with pseudomembranous colitis (Borriello and Barclay, 1985), and was shown to be highly virulent by comparative analysis of *C. difficile* strains in 1987 (Borriello *et al*, 1987). Extracellular toxin A levels were over 195ng/ml in Bactotryptose broth, while the maximum titre of toxin B was 1:40,960 following growth in tryptic nitrate broth, as detected in VERO cells . The strain was highly virulent in hamsters pre-treated with clindamycin, with all of the animals challenged dead or moribund within 48 hours (Borriello *et al*, 1987). B-1 cells were shown to be hydrophilic (Krishna *et al*, 1996), to produce a number of hydrolytic enzymes including chondroitin-4-sulphatase, hyaluronidase, heparinase and collagenase (Seddon *et al*, 1990), to show evidence of mucosal adherence (Borriello *et al*, 1988b), flagella and motility (Tasteyre *et al*, 2000a), and to posses capsule (shown by staining and electron microscopy) (Davies and Borriello, 1990), but not fimbrae (Borriello *et al*, 1988a).

The second strain, Tra5/5, was isolated in 1982 from the faeces of a healthy infant (Larson *et al*, 1982). It is of lower virulence (demonstrated by number of

animals killed after inoculation) with a significantly lower toxin A concentration than for strain B-1 (~25ng/ml), but the same maximum titre of toxin B, 1:40,960 following growth in tryptic nitrate broth, as detected in VERO cells (Borriello *et al*, 1987). Tra5/5 has also been shown to be hydrophilic (Krishna *et al*, 1996), and to produce the same range of hydrolytic enzymes as B-1 (chondroitin-4-sulphatase, hyaluronidase, heparinase and collagenase (Seddon *et al*, 1990). It shows evidence of fimbrae (Borriello *et al*, 1988a) and capsule formation (Davies and Borriello, 1990), but the mucosal adherence and presence of flagella has not been determined for this strain.

The third strain, designated A (donated by M. Wilcox, Leeds Teaching Hospitals), was isolated from the Stoke-Mandeville outbreak, associated with the emergence of the 'hypervirulent' ribotype 027 (Anonymous, 2005;Smith A., 2005). There have been fewer biochemical studies undertaken to characterise isolates of this outbreak, but it has been reported that these epidemic 027 strains produce higher levels compared to other ribotypes (toxin A, 848µg/L, toxin B 180µg/L (Warny *et al*, 2005)). They were shown to be motile (Stabler *et al*, 2009), and resistant to a number of antibiotics including fluoroquinolones, chloramphenicol and erythromycin (Stabler *et al*, 2009). The ribotypes of strains A, B-1 and Tra5/5have been analysed and confirmed as 027, 005 and 001 respectively in our laboratory.

2.1.2 – Growth Curves

Growth curves were determined in order to establish a time scale for optimum protein synthesis. Starter cultures in fastidious anaerobic broth (FAB) were prepared and used to incubate 25ml FAB to an optical density of 0.01-0.03 at 600nm. Absorbencies were read every 4 hours to produce a growth curve for each of the three strains.

2.1.3 – Growth conditions

Cells from each strain were taken from beads stored at -80ºC, plated onto Columbia blood agar (CBA) or fastidious anaerobe agar (FAA), and revived overnight at 37ºC in anaerobic conditions. Colonies from these plates were then used to inoculate fastidious anaerobe broth (FAB), or sub-cultured onto CBA or FAA to be used for protein extraction. Anaerobic conditions were maintained using anaerobic sachets (Oxoid, Basingstoke, UK).

2.1.4 – Extraction methods

There is no universal method for the extraction of proteins from microorganisms. Consequently, a number of different methods of protein extraction were compared using both broth and plate cultures. For each method the cell pellet was resuspended in lysis buffer (8M Urea, 2% CHAPS, 40mM Tris base) or 50mM phosphate buffer depending on downstream applications. The phenylmethanesulphonylfluoride (PMSF) protein inhibitor was added (0.5-1mM) and cells kept on ice throughout the process to reduce proteolysis.

i) - FastPrep method

Washed glass beads (≤ 106 μm) (Sigma-Aldrich, Dorset, UK) were added to approximately 1/3 sample volume, and samples were shaken on a *FastPrep* System (MP Biomedicals, Illkirch, France) for two runs of 60 seconds at 4m/s, and incubated on ice. For larger cell pellets, this was repeated up to five times, with five minute incubations on ice between each 2 x 60s run. Samples were centrifuged for 30 minutes at 21,000g, 4ºC, to remove cell debris. Protein extracts were stored at -20ºC.

ii) - Mickle Cell Disintegrator method

Washed glass beads (≤ 106 μm) (Sigma-Aldrich, Dorset, UK) were added to approximately 1/3 sample volume and samples were disrupted using the *Mickle Cell Disintegrator* (Mickle Laboratory Engineering, Guidford, UK) for 15 minutes before incubation at -80ºC for 15 minutes. This was repeated, and samples were then centrifuged for 30 minutes at 21,000g, 4ºC, to remove cell debris, and protein extracts stored at -20ºC.

iii) - French press method

Resuspended samples were passed through the *French Press* (Thermo Fisher Scientific, San Jose, USA) three times, then centrifuged for 30 minutes at 21,000g, 4ºC to remove cell debris. Protein extracts were stored at -20ºC.

2.1.5 - 1D gel electrophoresis

Single dimension gel electrophoresis was carried out using the NuPAGE® gel system (Invitrogen, Paisley, UK). NuPAGE® LDS sample buffer (4μl) and

NuPAGE® sample reducing agent (1.6µl) was added to 5µg protein sample and made up to 16µl with distilled water. Each sample was then loaded onto a sample well of a NuPAGE® 10% (w/v) bis tris gels alongside 1µg Protein Molecular Weight Standards (P-6649 Molecular Probes, Invitrogen, Paisley, UK). Gels were run in NuPAGE® MES SDS running buffer for 35 minutes, 200V, 120mA, 25W. Proteins were fixed in 10% (v/v) glacial acetic acid, 40% (v/v) methanol for 1 hour prior to staining with SYPRO® Ruby Protein Gel stain (BioRad, Hemel Hempstead, UK) overnight in the dark. Gels were destained in 7% (v/v) acetic acid, 10% (v/v) methanol for 30 minutes before scanning by the Ettan DIGE imager (GE Healthcare, Chalfont St Giles, UK). Gels were post-stained with Brilliant Blue G-colloidal (Sigma, St Louis, USA) (0.1% (w/v) Brilliant Blue G, 0.29 M phosphoric acid and 16% saturated ammonium sulphate) according to the manufacturer's protocol.

2.2 – Protein Identification

2.2.1 – Protein Identification I – 2D Reference mapping

2.2.1.1 - Growth conditions for standard protein extracts

All strains were grown on Columbia blood agar (CBA) at 37ºC. Anaerobic conditions were maintained using anaerobic sachets (Oxoid, Basingstoke, UK). Cells were harvested for protein extraction after 64 hours, when the majority of cells were estimated to be in stationary phase, and optimum reproducible protein profiles generated.

2.2.1.2 - Preparation of standard whole cell protein extract

The growth from CBA plates was harvested after 64 hours. Cells were resuspended in lysis buffer (8M Urea, 2% CHAPS, 40mM Tris base) containing PMSF protease inhibitor (0.5-1mM) for cell lysis and protein extraction. Washed glass beads (≤ 106 μm) (Sigma-Aldrich, Dorset, UK) were added to the cell suspension, and homogenisation was performed on the *FastPrep* System (MP Biomedicals, Illkirch, France). Samples were shaken for 2 runs of 60 seconds at 4m/s, and incubated on ice before centrifugation for 30 minutes at 21,000g, 4ºC, to remove cell debris. The protein concentration of the lysates was determined using Bradford's assay, where protein lysates were diluted 1:20 and incubated 1:50 with Bradford's reagent for 10 – 30 minutes. Absorbance was read at 630nm and compared to a standard bovine serum albumin (BSA) curve to estimate the protein concentration. Protein extracts were stored at -20ºC.

2.2.1.3 - 2D gel electrophoresis

After being treated with the GE 2D clean up kit (GE Healthcare, Chalfont St Giles, UK) according to the manufacturer's protocol, proteins were resuspended in DIGE lysis buffer (30mM Tris, 2M Thiourea, 7M Urea 4% (w/v) CHAPS). 50-100μg protein was diluted to 100μl with rehydration buffer (7 M Urea, 2 M Thiourea, 2% (w/v) CHAPS, 0.5%(v/v) IPG Buffer, 0.002% (v/v) bromophenol blue, 2.8mg/ml DTT) and loaded onto hydrated immobiline dry strips (18cm) for isoelectric focussing. These were run overnight before equilibration with DTT (6 M Urea, 75mM Tris-HCl pH 8.8, 29.3% (v/v) glycerol, 2% (w/v) SDS, 0.002% (v/v) bromophenol blue, 65mM DTT) for 15 minutes followed by

Iodoacetamide (6 M Urea, 75mM Tris-HCl pH 8.8, 29.3%(v/v) glycerol, 2% (w/v) SDS, 0.002%(v/v) bromophenol blue, 135mM Iodoacetamide) for 15 minutes. The second dimension was run on 10% acrylamide gels (22 x 24 cm x 1 mm) for 6 hours. Proteins were fixed in 10% (v/v) glacial acetic acid, 40% (v/v) methanol for 1 hour prior to staining with SYPRO ruby Protein Gel stain overnight in the dark. Gels were destained in 7% (v/v) acetic acid, 10% (v/v) methanol for 30 minutes before being scanned using the Ettan DIGE imager (GE Healthcare, Chalfont St Giles, UK). Gels were then post-stained with Brilliant Blue G-colloidal (Sigma-Aldrich, Gillingham, UK) (0.1% (w/v) Brilliant Blue G, 0.29 M phosphoric acid and 16% saturated ammonium sulphate) so that spots could be visualised for manual picking.

2.2.1.4 - In-gel digestion

Spots from 2D gels were destained in 25mM ammonium bicarbonate, in 50% (v/v) methanol, until no blue colour remained. Gel plugs were then dehydrated with two 10 minute incubations in 100% acetonitrile, and air dried for 10 minutes before reduction in 10mM DTT in 25mM ammonium bicarbonate (30 minutes at 60ºC), and alkylation in 55mM IAA in 25mM ammonium bicarbonate, (45min in the dark). The plugs were washed three times in 25mM ammonium bicarbonate and dehydrated as above, before the digestion step. 10µl trypsin in 25mM ammonium bicarbonate solution (10ng/µl) was added to each gel plug, and the plate was incubated at 37ºC for 16 hours overnight. Peptides were extracted in 0.05% (v/v) trifluoroacetic acid (TFA), 50% (v/v) acetonitrile (10µl) for direct spotting onto MALDI plates for MALDI-TOF analysis, or 0.1% (v/v) TFA for LC-MS analysis.

2.2.1.5 - MALDI-TOF mass spectrometry analysis

After tryptic digestion, 0.7µl of each peptide solution was co-crystallised on the target plate with 0.7µl matrix solution (10mg/ml α-cyano-4-hydroxycinnamic acid in 49.5% (v/v) acetonitrile, 49.5% (v/v) ethanol, 0.001% (v/v) TFA). Peptides masses were measured on a MALDI-TOF *Reflectron* (Waters, Milford, USA) equipped with a 337nm nitrogen laser. Analysis was performed in positive ion mode. Peptide masses were collected over a m/z range of 800-3000 Da, and 1.8µg/ml rennin in 0.1% (v/v) TFA was used to lock mass the instrument. Raw data files were then compared with protein sequence data from the National Centre for Biotechnology Information (NCBI) database, and a redundant database comprising strains A, B-1,Tra5/5 and 630, using the online Mascot Wizard software (http://www.matrixscience.com/wizard.html). One missed cleavage per peptide was allowed, and a mass tolerance of 150 ppm was used. Carbamydomethylation of cysteine was specified as a fixed modification, and oxidation of methionine as a variable modification. A MASCOT score greater than 54 was significant ($p<0.05$), but protein identifications with scores lower than 54 were considered positive if they also showed a minimum of 6 matched peptides or a sequence coverage of over 25% (as described previously (Encheva *et al*, 2009)).

2.2.2 – Protein identification II – LC-MS/MS

2.2.2.1 - 1D gel electrophoresis

As above (2.1.5)

2.2.2.2 - In-gel digestion

1D gel lanes were cut into 12 bands and each band was destained and digested as above. Peptides were extracted in 0.1% (v/v) TFA for LC-MS analysis.

2.2.2.3 - LC-MS/MS mass spectrometry analysis

The peptide mixtures from tryptic digestion of gel bands from 1D PAGE were separated and analysed using an Ultimate 3000 Dionex nano/capillary HPLC system (Dionex (UK) Ltd., Camberley, UK) coupled with a Thermo LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, USA).

Trypsin-digested peptides (5 μl) were cleaned and desalted on a reversed phase trap column (PepMap C18, 300 μm i.d. x 5 mm, 3μm, 100 Å, (Dionex (UK) Ltd., Camberley, UK)), separated on a nano C18 analytical column (PepMapC18, 75 μm i.d. x 15 cm, 3μm, 100 Å, (Dionex (UK) Ltd., Camberley, UK)), followed by analysis of separated peptides with electrospray tandem mass spectrometry (MS/MS) on the LTQ Orbitrap mass spectrometer.

Settings for the LC-MS/MS were as follows: peptide mixtures were initially trapped on the reversed phase trap column using isocratic of 100% solvent A (2% acetonitrile, 0.1% formic acid in water) at a flow rate of 25 μl/min. Subsequent separation was performed on the analytical C18 nano column using a 45-minute gradient of 5 to 45% solvent B (90% acetonitrile, 0.1% formic acid in water) versus solvent A, then to 90% B for an additional 5 minutes with a flow rate of 300nl/min. The mass spectrometer was operated in a data-dependent mode to automatically switch between MS and MS/MS acquisition. The full survey scan (m/z 440–2000) was acquired in the Orbitrap with a resolution of

60,000 at *m*/*z* 400, which was followed by six MS/MS scans in which the most abundant peptide precursor ions detected in the preceding survey scan were dynamically selected and subjected for collision-induced dissociation (CID) to generate MS/MS spectra, which were later searched against appropriate databases to assign identities to proteins.

The electrospray voltage was set at 1.5 kV, the ion source capillary voltage and temperature were set at 35 V and 200ºC, respectively. Tube lens was 105 V and the normalised collision energy was 35% for MS/MS. Enhanced mass accuracy was achieved using a background polydimethycyclosiloxane ion with m/z = 445.120025 was used as 'lock mass'. Real time recalibration on the 'lock mass' by corrections of mass shift removed mass error associated with calibration of the mass scale. A mass accuracy of 1-5 parts per million (ppm) was routinely maintained during data acquisition.

The raw data files were matched against a theoretical tryptic digest of the resulting proteins of a redundant *C. difficile* database containing genomic data from strains A, B-1 and Tra 5/5 and the 630 reference strain using the MASCOT daemon software (Matrix Science, London, UK), with a tolerance of two missed cleavages. A peptide mass tolerance of 10 PPM for the parent ion, and 0.50 Da for the fragment ions was specified, allowing for carbamidomethylation of cysteine as a fixed modification, and oxidation of methionine as a variable modification. The software program, Scaffold (Proteome software, Portland, USA) was used to verify and compare protein identifications between strains. A minimum of two peptides was required for protein identification, although

identifications based on only one peptide were considered valid if the detected peptide was present in the translated genome of the relevant strain. Identifications of non-annotated hypothetical proteins were given putative functions by finding similar proteins in the NCBI database BLAST. Where possible, identifications were matched to the equivalent protein identification from the 630 reference strains, allowing redundant proteins to be identified and removed.

2.2.2.4 - BLAST searching of protein sequences against the unfinished genomes of each strain

The sequence of each identified protein was searched against each unfinished genome using the pBLAST stand alone protein BLAST search, with tab deliminated outputs and an e-value threshold of 10^{-5}. Default settings were used for all other parameters. The percentage identification and e-values for each protein against each genome was recorded, and a percentage identification greater than 75% was taken to indicate the presence of the corresponding gene in a strain. Protein sequences were also BLAST searched against the 630 reference strain genome.

2.2.3. – Protein Identification III – DIGE analysis

2.2.3.1 – CyDye labelling of protein samples

Protein samples were cleaned using the GE 2D clean up kit according to the manufacturer's protocol and resuspended in DIGE lysis buffer. Protein samples were diluted to 5µg/µl and the pH adjusted to 8.0 if necessary, before

minimal labelling with 400pmol CyDye per 50µg protein (CyDye DIGE Fluor, GE Healthcare, Chalfont St Giles, UK). CyDye was reconstituted according to the manufacturer's protocol with 5µl Dimethyl formamide (DMF)(Aldrich, anhydrous, 99.8% pure, Sigma-Aldrich, Gilliangham, UK), to give a 1mM stock solution. CyDye stock was mixed with DMF (1:1.5) to give 400µM labelling solution. 400pmol of CyDye (1µl labelling solution) was added per 50µl protein sample, vortexed briefly, spun down and incubated on ice for 30 minutes in the dark. 1µl 10mM lysine was added per 50µl protein sample to stop the reaction and the samples incubated on ice in the dark for a further 10 minutes before use or storage.

2.2.3.2 – 2D DIGE

The CyDye labelled protein samples were diluted with an equal volume of rehydration buffer (7 M Urea, 2 M Thiourea, 2% (w/v) CHAPS, 0.5% (v/v) IPG Buffer, 0.002% (v/v) bromophenol blue, 2.8mg/ml DTT) before being pooled, and the pooled samples then run on the first and second dimensions as described above (2.2.1.3).

2.2.3.3 – Experimental setup

Two biological replicates for each of the three strains were used, giving six samples in total. An internal standard of a mix of all six samples was run on each gel, alongside two individual samples. Gels were run for each strain against the two other strains, and repeated with the biological replicates to give a total of four profiles for each strain. Each biological replicate was labelled with both Cy3 and Cy5 during the experiment. Spot detection, spot alignment,

and statistical analysis of the relative protein concentrations between the strains was carried out using the Progenesis SameSpots software (Nonlinear Dynamics, Newcastle, UK).

2.3 - ProteoMiner Treatment

2.3.1 - Growth Conditions

Large amounts of protein were required for ProteoMiner treatment, so a much larger-scale approach was adopted. The three strains, A, B-1 and Tra5/5, were grown in batches of 3 x 500ml Fastidious Anaerobic broth (FAB) for 24 hours at 37ºC. Anaerobic conditions were maintained using anaerobic sachets (Oxoid, Basingstoke, UK). As controls, 100µl culture was plated onto Columbia blood agar plates and incubated under anaerobic or aerobic conditions. If anaerobic controls were positive and aerobic controls were negative, cells were harvested by centrifugation at 8,000g, 4ºC, for 30 minutes.

2.3.2 - Protein extraction methods and comparison

Cells were washed three times in 1x TE buffer or until the extrapolysaccharide layers were depleted, then resuspended in 50mM phosphate buffer containing 0.5mM PMSF for cell lysis and protein extraction. Glass beads were added to the cell suspension, and homogenisation was carried out on the *FastPrep* System (MP Biomedicals, Illkirch, France). Samples were shaken for 60 seconds at 4m/s, and incubated on ice for five minutes. This was repeated five times before centrifugation for 30 minutes at 21,000g, 4ºC, to remove cell debris. The protein

concentration of the lysates was determined using Bradford's assay. Protein extracts were stored at -20ºC. For each strain, it was necessary to pool several extractions in order to generate sufficient protein for ProteoMiner treatment.

2.3.3 - ProteoMiner treatment

The pooled protein extracts were divided into two technical controls, and each 15ml protein extract was treated with an 'E-bead' amino terminus hexapeptide library (250µl), followed by an 'SE-bead' carboxylated library (280µl). Dry beads were rehydrated by incubation with dH_2O containing approximately 5% EtOH for 5 minutes with end-to-end rotation, then centrifuged for 1 minute at 4000 rmp, and the supernatant discarded. The beads were then equilibrated by washing in 20x PBS buffer (5 minutes, pH 6-8), followed by 5 minute washes in dH_2O and 1x PBS. After equilibration, an equal volume of 1x PBS was added to the bead to give a 50% slurry, which was distributed into 15ml tubes to give the required volume of beads in each tube, before the beads allow to settle and the supernatant discarded.

Protein extracts were added to the E-bead hexapaptide libraries, and incubated for 1.5 hours, with rotation, at 4ºC. The supernatant was recovered by centrifugation at 4,000g for 1 minute, added to the SE-bead hexapeptide library, and incubated as before. Again, the supernatant was removed, and the two hexapeptide bead libraries with bound proteins were resuspended in 1x PBS, transferred to spin columns and washed a further 5 times with 1x PBS. Waste PBS was removed by centrifugation at 4,000g for 1 minute, and the beads incubated with elution buffer (8M Urea, 2% CHAPS, 50mM citric acid) for 10

minutes to dissociate bound proteins from the hexapaptide bead libraries. The eluant was collected by centrifugation at 4,000g for 1 minute. The elution process was repeated a further two times, and the three eluant fractions from each library pooled.

The total eluant was halved, and the first fraction frozen for 2D clean up. The second fraction was dialysed against 0.1M ammonium bicarbonate for LC-MS and further analysis.

2.2.4 - Dialysis of treated samples

Samples were injected into dialysis membranes (Slide-A-Lyzer 3.5K dialysis cassettes (0.5 - 3.0 ml sample volume), Pierce, Rockford, USA) and incubated in a 2L beaker of 0.1M ammonium bicarbonate overnight at 4ºC. The buffer solution was then changed and the membranes incubated for a further two hours at 4ºC before the sample was removed, aliquoted and stored at -20 ºC.

2.3.5 - 2D gel electrophoresis

After being treated with the GE 2D clean up kit according to the manufacturer's protocol, 50-100µg protein was diluted to 100µl with rehydration buffer and run on the first and second dimension as described above. Gels were post-stained with Sigma Brilliant Blue G-colloidal so that spots could be visualised for manual picking. In gel digestion, MALDI-FOF MS and 1D GE followed by LC-MS/MS were all carried out as described above (2.2.1.3).

2.4 - Western Blotting

Proteins were separated by 1D or 2D electrophoresis as described above.

2.4.1 - Protein transfer

Proteins were not fixed or stained in the gels used for Western blotting. Gels were imaged after running, and then equilibrated in Towbin transfer buffer (25mM Tris, 192mM Glycine, 0.1% SDS, 20% methanol) for five minutes. For each gel, 4 sheets of blotting paper and 1 of hybond nitrocellulose membrane (GE Healthcare, Chalfont St Giles, UK) were cut to the size of the gel. For 1D gels (6x8cm) approximately 10ml Towbin transfer buffer per sheet was used, and for 2D gels (20x20cm) approximately 50ml Towbin transfer buffer per sheet was used. The nitrocellulose membrane was pre-wetted with distilled water, then soaked in Towbin transfer buffer for 5 minutes.

Protein transfer was carried out on a TE77 semi-dry blotting unit (Invitrogen, Paisley, UK). Transfer from 1D gels was carried out at 45mA for 1 hour, and transfer from 2D gels was carried out at 320mA for 1 hour.

After transfer, membranes were stained or imaged to ensure transfer was complete.

2.4.2 - CyDye labelling

CyDye labelling was carried out as described for the DIGE analysis above.

2.4.3 - Staining

If protein samples had been labelled with CyDye before electrophoresis, gels could be imaged before transfer and membranes imaged after transfer without the need for staining. If protein samples were unlabelled, two gels were run in parallel, one gel fixed, stained and imaged as described above, and the other used for transfer. After transfer the membrane could be imaged using either Ponceau S solution or Sypro Ruby blot stain.

i) Ponceau S solution

After blotting, the membrane was stained with Ponceau S solution (0.1% in 5% Acetic Acid) for approximately 10 minutes until proteins become visible. The membrane was the destained in 5% TCA before blocking and probing.

ii) Sypro Ruby Blot strain

After blotting, the membrane was immersed in 7% acetic acid, 10% Methanol for 15 minutes with shaking, before being rinsed in distilled water for 5 minutes, four times. The membrane was incubated in Sypro Ruby blot stain (Invitrogen, Paisley, UK) for 15 minutes in the dark, then washed 4-6 times in distilled water to remove excess dye before imaging.

2.4.4 - Probing of Western blots

After transfer, membranes were rinsed in 0.01M PBS/Tween (0.138M NaCl, 0.0027M KCl, 0.05% Tween 20, (Sigma)) before being blocked with 1% BSA/PBS (0.138 M NaCl, 0.0027M KCl, 1%(w/v) BSA, (Sigma-Aldrich, Gillingham,UK)). Between incubation steps, membranes were washed in PBS/Tween for three

times 5 minutes. Once conditions had been optimised, antisera were used at a dilution of 1:10,000 in 0.5% BSA, 0.01M PBS, and primary antibodies detected using cy5 labelled anti-rabbit IgG (AbCam, Cambridge, UK) or anti-human IgG (AbCam, Cambridge, UK) at a dilution of 1:10,000 in 0.5% BSA, 0.01M PBS). All incubations were carried out for at least 3 hours at room temperature, and between incubations membranes were washed in PBS/Tween for three times 5 minutes.

The small (70mm x 90mm) membranes from 1D gels were incubated in 10ml antibody solution in plastic incubation boxes, giving good detection of immunoreactive proteins. However, incubating large 200mm x 200mm membranes in plastic 2D gel boxes required approximately 30 ml antibody solution to cover the membrane, so lower volume alternatives were considered. Hybridisation tubes were trialled. 130mm x75mm Hybritube-20 (GIBCO-BRL, Invitrogen, Paisley, UK) tubes were trialled vertically, and 240mm x 120mm mm HYBAID tubes were trialled horizontally in a rotating hybridisation oven (Thermo-Fisher Scientific, San Jose, USA). This required only 13 or 15 ml antibody solution respectively, and improved the detection of immunoreactive proteins. However, the diameter of the hybridization tubes meant that the membrane within the tube overlapped, which could affect the access of the antibody to the bound proteins. Therefore hybridization bags (240mm x 250mm) were used. 15ml antibody solution was required, and the bags incubated on a circular tube rotator.
Membranes were imaged on the Ettan DIGE imager (GE Healthcare, Chalfont St Giles, UK).

2.4.5 – Antibodies used

i) Primary antibodies

- Rabbit polyclonal to horse serum (H8890, Sigma, Gillingham, UK)
- Rabbit serum (N. Fairweather, Imperial College, London) - antibodies raised against *Clostridium difficile* strain 630 by injecting rabbits with a heat killed preparation of whole bacterial cells.
- Human serum (I.R. Poxton, University of Edinburgh) - pooled serum samples from patients;

A) Cases – symptomatic, toxin-positive and culture-positive
B) Carriers – asymptomatic, culture positive, toxin-variable
C) Controls – asymptomatic, toxin-negative, culture-negative

Controls were matched for age, and were in the same wards as patients in groups A and B.

ii) Secondary antibodies

- Cy3 labelled goat polyclonal to Rabbit IgG (Ab6939, AbCam, Cambridge, UK)
- Cy5 labelled goat polyclonal to Rabbit IgG (Ab6564, AbCam, Cambridge, UK)

- Cy3 labelled goat polyclonal to Human IgG (Ab6957, AbCam, Cambridge, UK)
- Cy5 labelled goat polyclonal to Human IgG (Ab6561, AbCam, Cambridge, UK)

Chapter 3 – Results

3.1 – Growth Curves and Optimisation of Protein Extraction

3.1.1 - Growth Curves

In order to confirm that the three strains progress through the growth cycle at a similar rate, growth curves were derived and compared. To obtain a complete curve, but not interrupt anaerobic conditions too often, replicates were used and optical densities measures at different time intervals. Readings were pooled and averaged to give the growth curve shown in *Figure 3.1*.

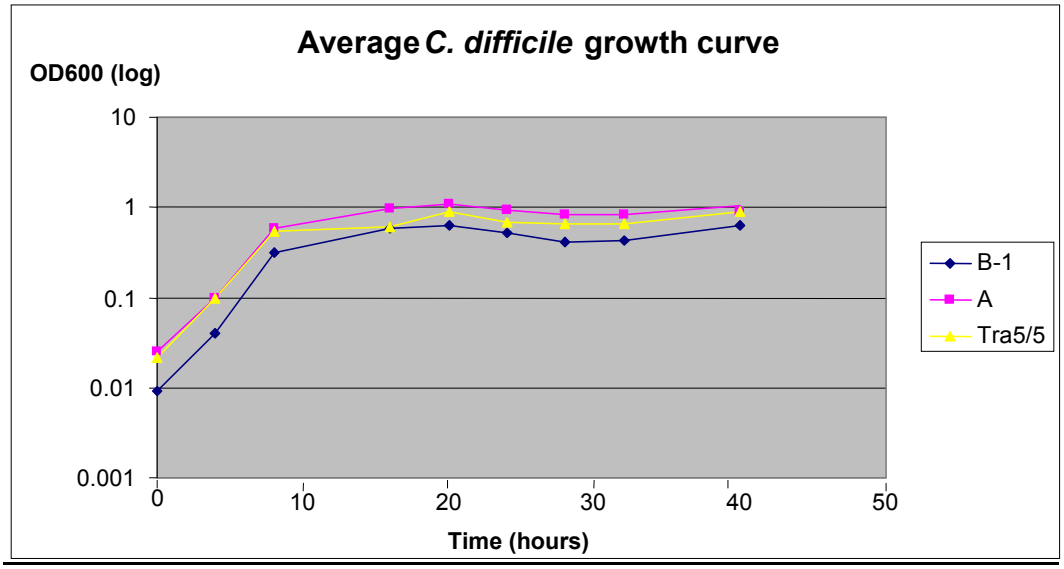


Figure 3.1 *– A growth curve comparing growth rates of the three different strains of Clostridium difficile, A, B-1 and Tra5/5. Strains were grown in Fastidious Anaerobic Broth (FAB) and optical densities at 600nm taken at a number of different time intervals. Six replicates were used for each strain, and the time points the optical densities were read were varied for different sets of replicates to get a range of*

readings across the growth curve. Readings for all replicates were combined and averaged.

The three strains were found to follow similar growth curves and enter stationary phase after 14 hours (*Figure 3.1)*, so cultures of different strains grown for the same length of time are likely to be in the same phase of growth. This minimises the possibility that any differences observed in protein concentration are due to phase expression variations, and allows any strain-to-strain expression differences to be detected.

3.1.2 - Extraction method

A number of different methods are available for extracting proteins from bacterial cultures. In order to determine the most straightforward and reproducible extraction protocol, three different methods were compared, The French Press (Thermo Fisher Scientific), the Mickle beater (Mickle Engineering) and the Fast Prep (MP Biomedicals). Proteins were extracted from cultures grown on CBA plates for 24 hours.

Insufficient protein was generated by the French Press method for downstream application due to the absence of a micro pressure cell, so this method was discounted. Comparable amounts of protein were obtained using Mickle and Fast Prep methodologies, and the protein profiles of the extracts were similar (*Figure 3.2*). The Fast-Prep method was the quickest and most straightforward, and so was used in all subsequent extractions.

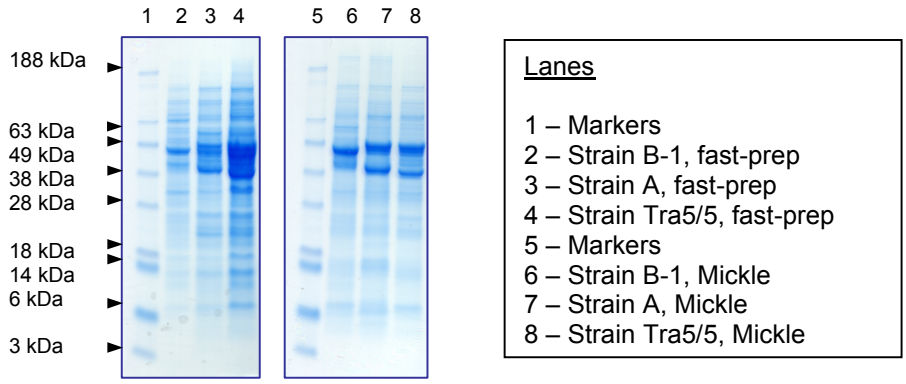


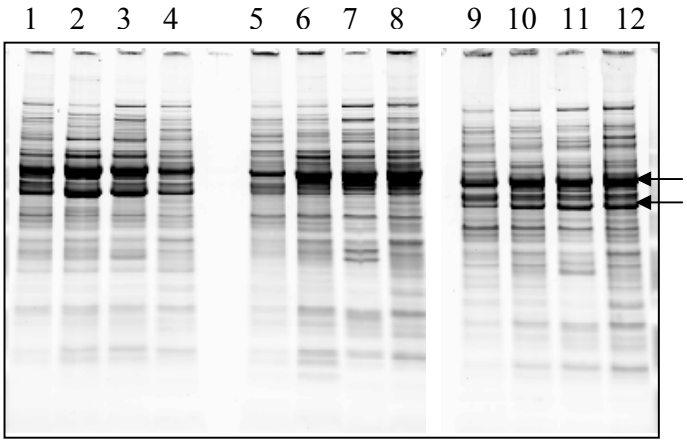
Figure 3.2 *– Comparison of Coomassie Brilliant Blue stained 1D protein profiles of Mickle and Fast Prep extractions. Strains were grown on Columbia Blood Agar (CBA) and protein extractions are shown alongside a molecular ladder for sizing.*

3.1.3 - Growth conditions

For comparison of expression levels between strains, a standard protein extract with a reproducible and extensive profile was required. To determine the best growth conditions to achieve this, whole cell protein extractions from a number of different cultures grown for different lengths of time, and in different media, were compared.

Firstly, extractions from broth cultures at different points on the growth curve were compared (*Figure 3.3)*.

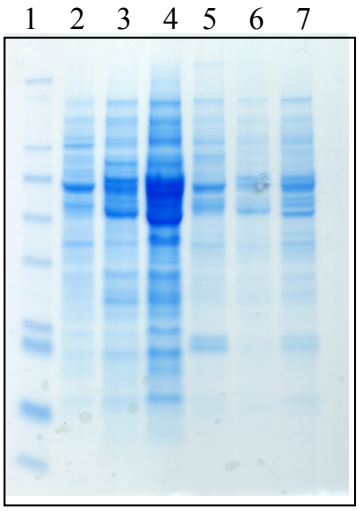
The protein profiles of cultures grown for different lengths of time do not vary enormously, but additional proteins are clearly visible in extracts from cultures grown for 36 hours, when according to the growth curves, cultures are in stationary phase. The most abundant proteins of ~35 and 45 kDa are likely to correspond to the S-layer proteins.



Lanes

1 – Strain B-1, 6 hours
2 – Strain B-1, 16 hours
3 – Strain B-1, 24 hours
4 – Strain B-1, 36 hours
5 – Srain A, 6 hours
6 – Strain A, 16 hours
7 – Strain A, 24 hours
8 – Strain A, 36 hours
9 – Srain Tra5/5, 6 hours
10 – Strain Tra5/5, 16 hours
11 – Strain Tra5/5, 24 hours
12 - Strain Tra5/5, 36 hours

Figure 3.3 *– Comparison of Sypro Ruby stained 1D protein profiles of extractions from broth (FAB) cultures grown for different lengths of time. Arrows indicate the position of probable S-layer proteins.*



Lanes

1 – Markers
2 – Strain B-1, plates
3 – Strain A, plates
4 – Strain Tra5/5, plates
5 – Strain B-1, broth
6 – Strain A, broth
7 – Strain Tra5/5, broth

Figure 3.4 *– Comparison of Coomassie Brilliant Blue stained 1D protein profiles of extractions from plate cultures (CBA) and broth cultures (FAB) incubated for 24 hours.*

Profiles of extracts from broth cultures and plate cultures were compared (*Figure 3.4*). The most notable difference in the plate and broth extracts is that the S-layer proteins are less abundant in the broth extracts than the plate extracts. However, a greater number of proteins are visible in the plate extracts. Extraction from plates is simple and straightforward, and allows the presence of contaminants to be detected easily. It does not require the centrifugation and washing steps necessary in extraction from broth culture, so fewer cells are lost. As the extracts from plate culture appeared to give more extensive profiles, this was chosen for the standard extracts.

Cultures grown on Columbia Blood Agar (CBA) and Fastidious Anaerobe Agar (FAA) were compared, and found to give comparable profiles (*Figure 3.5)*. More proteins were visible in the extracts grown for 64 hours (*Figure 3.5*), so these were chosen for the standard extractions.

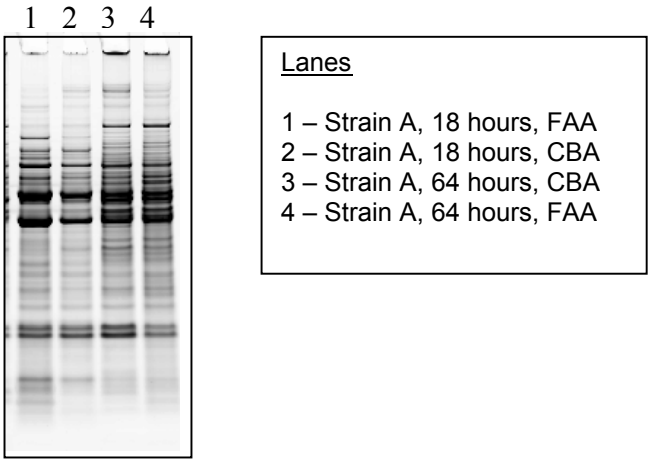


Figure 3.5 *– Comparison of Sypro Ruby stained 1D protein profiles of extractions from strain A cultures grown on CBA and FAA for 18 and 64 hours.*

3.2 – Protein Identification

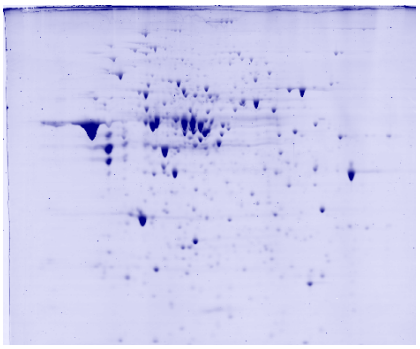
3.2.1 – Protein Identification I – 2D Reference Mapping

The standard protein extracts from plates were compared to extractions from broth culture by 2D gel electrophoresis (*Figure 3.6)* to confirm that the plate extractions gave reproducible and extensive 2D, as well as 1D, profiles. A large number of protein spots, and good resolution, was obtained for both extraction methods, so the standard plate extractions were used for reference mapping.

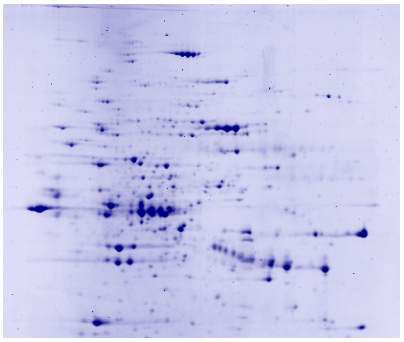
The 2D protein profiles of the three strains were similar, showing consistent size and pI range, and comparable number of protein spots.

To determine the 2D conditions which give the best spot resolution for mapping and subsequent MALDI-TOF-MS, gels were run using a variety of first dimension pH ranges (*Figure 3.7*). Very few basic proteins were resolved on any gels, and the 4-7 pH range gave the best resolution of the greatest number of proteins, so this pH range was used for the reference maps.

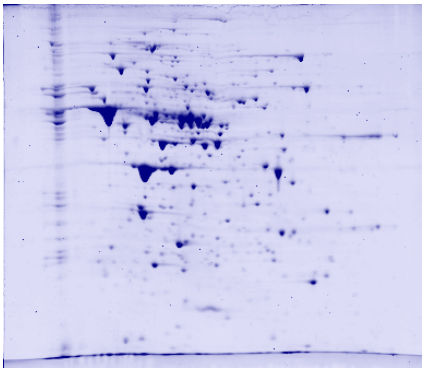
i) - *Strain B-1 protein extract from broth culture*



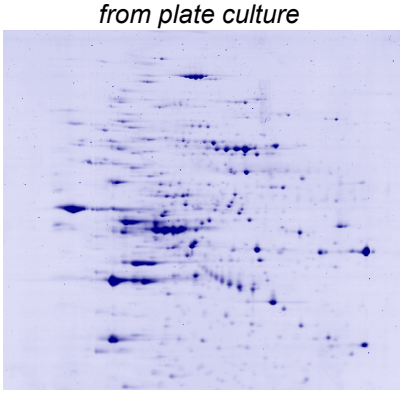
ii) - *Strain B-1 protein extract from plate culture*



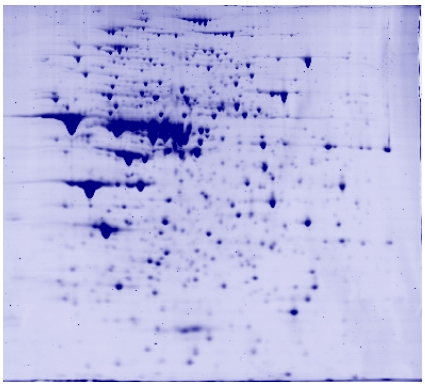
iii) - *Strain A protein extract from broth culture*



iv) - *Strain A protein extract from plate culture*



v) - *Strain Tra5/5 protein extract from broth culture*



vi) - *Strain Tra5/5 protein extract from plate culture*

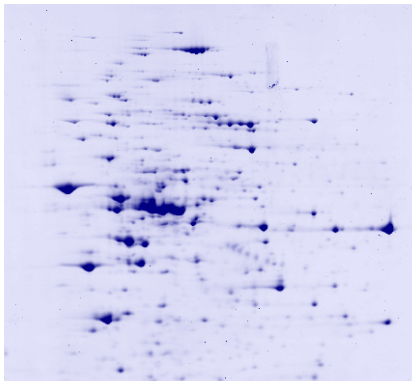
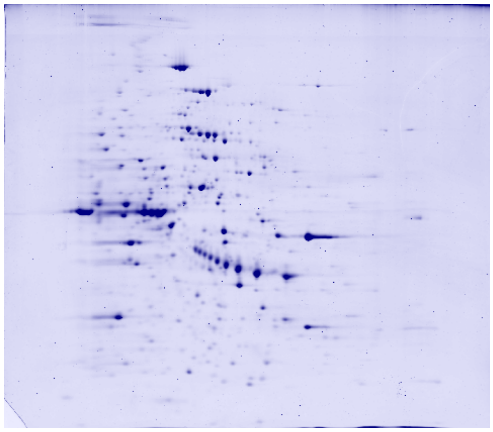
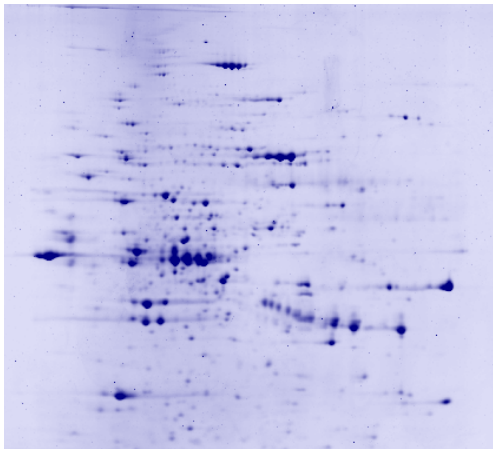


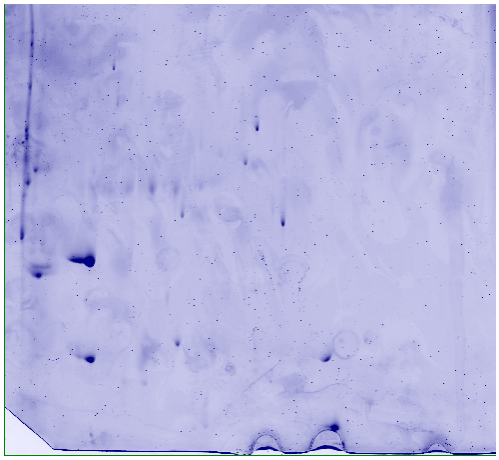
Figure 3.6 *- Comparison of Sypro Ruby stained 2D gel profiles of standard plate (CBA, 64h) and broth (FAB, 24h) extractions for strains A, B-1 and Tra5/5. Samples were run in triplicate, but only one replicate for each sample is shown here. In all cases profiles of proteins expressed differed between growth conditions, but many protein spots were visualised in both extracts.*



i)– *pH range 3-10*



ii) – *pH range 4-7*



iii) – *pH range 6-11*

Figure 3.7 *- Comparisons of the strain B-1 standard protein extract separated by 2D gel electrophoresis using different pH ranges for the first dimension. The sample was run in duplicate at each pH range, but only one replicate is shown here. The 4-7 pH range shown in ii) gives the best spot resolution*

One picking gel was chosen per strain for creation of the reference maps (*Figures 3.8, 3.9, 3.10*). Protein spots were excised manually before being subjected to in-gel digestion with trypsin and identification by MALDI-TOF MS. Identifications are shown with MASCOT scores, peptide coverage and number of matched peptides in Appendix Table A1. A MASCOT score greater than 54 was significant ($p<0.05$), but protein identification with scores lower than 54 were considered positive if they also showed a minimum of six matched peptides or a sequence coverage of over 25%.

For strain A, 192 spots were picked from the gel (*Figure 3.8*), of which 107 spots (56%) were identified by MALDI-TOF MS (*Appendix A1)*. Seven spots consisted of two or more different proteins and a total of 80 different proteins were identified on the strain A reference map.

One hundred and sixty eight spots were picked from the gel for strain B-1 (*Figure 3.9*), of which, 126 (75%) were identified by MALDI-TOF MS (A*ppendix A1)*. Seven of these consisted of two or more different proteins. A total of 95 different proteins were identified on the strain B-1 reference map.

For strain Tra5/5, 192 spots were picked (*Figure 3.10*). Of these, 158 (82%) were identified by MALDI-TOF MS (*Appendix A1*), and seven consisted of two or more different proteins. A total of 115 different proteins were identified for strain Tra5/5.

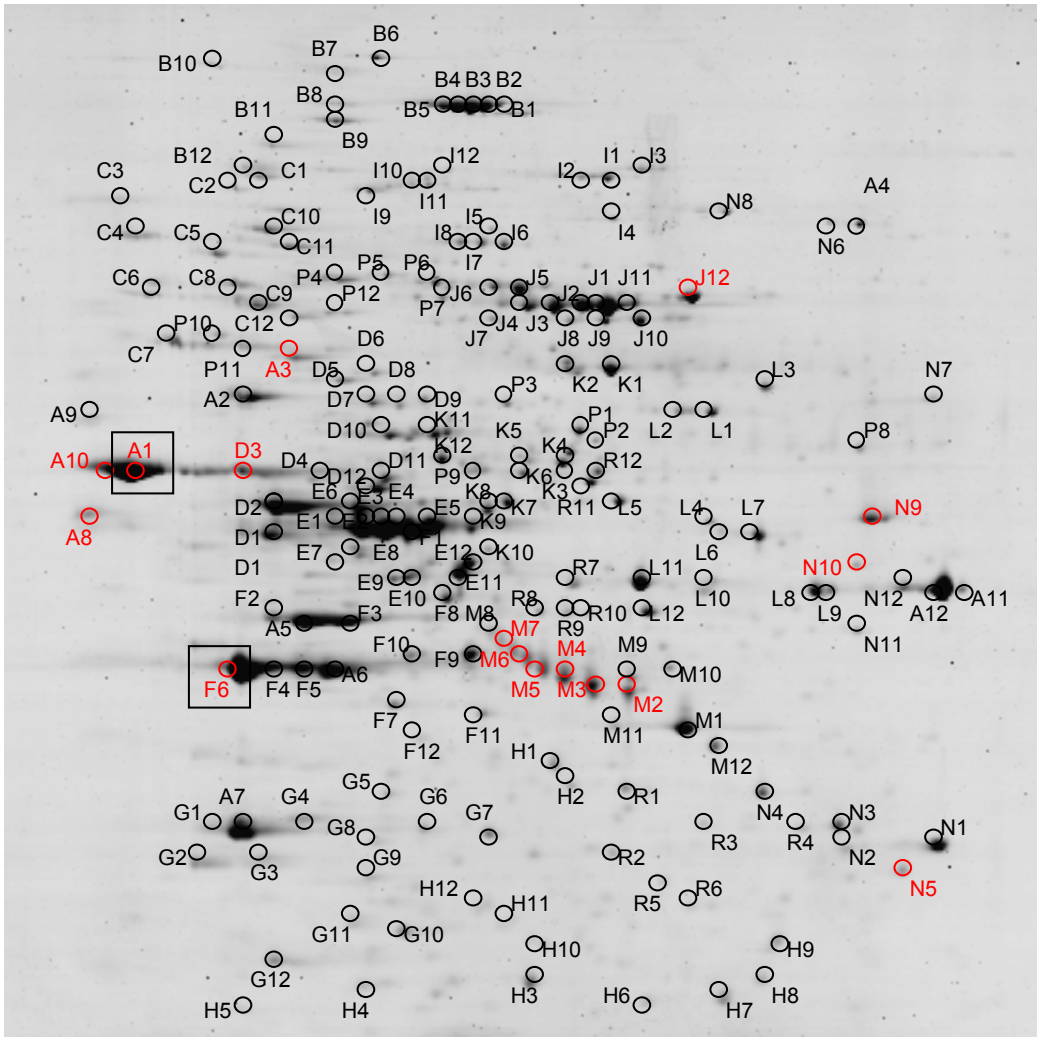


Figure 3.8 *– Spots picked for in-gel digestion and identification on the strain A reference map. Identifications are shown in Appendix A1. Each spot was assigned an identity letter, A-P, and number 1-12 as they were excised. To a certain extent, spots in the same position on different maps were given the same identity number, but variation between gels means this is not always the case. Surface and virulence proteins are highlighted in red, and the S-layer proteins corresponding to the SlpA gene product of the 630 reference strain are highlighted by boxes.*

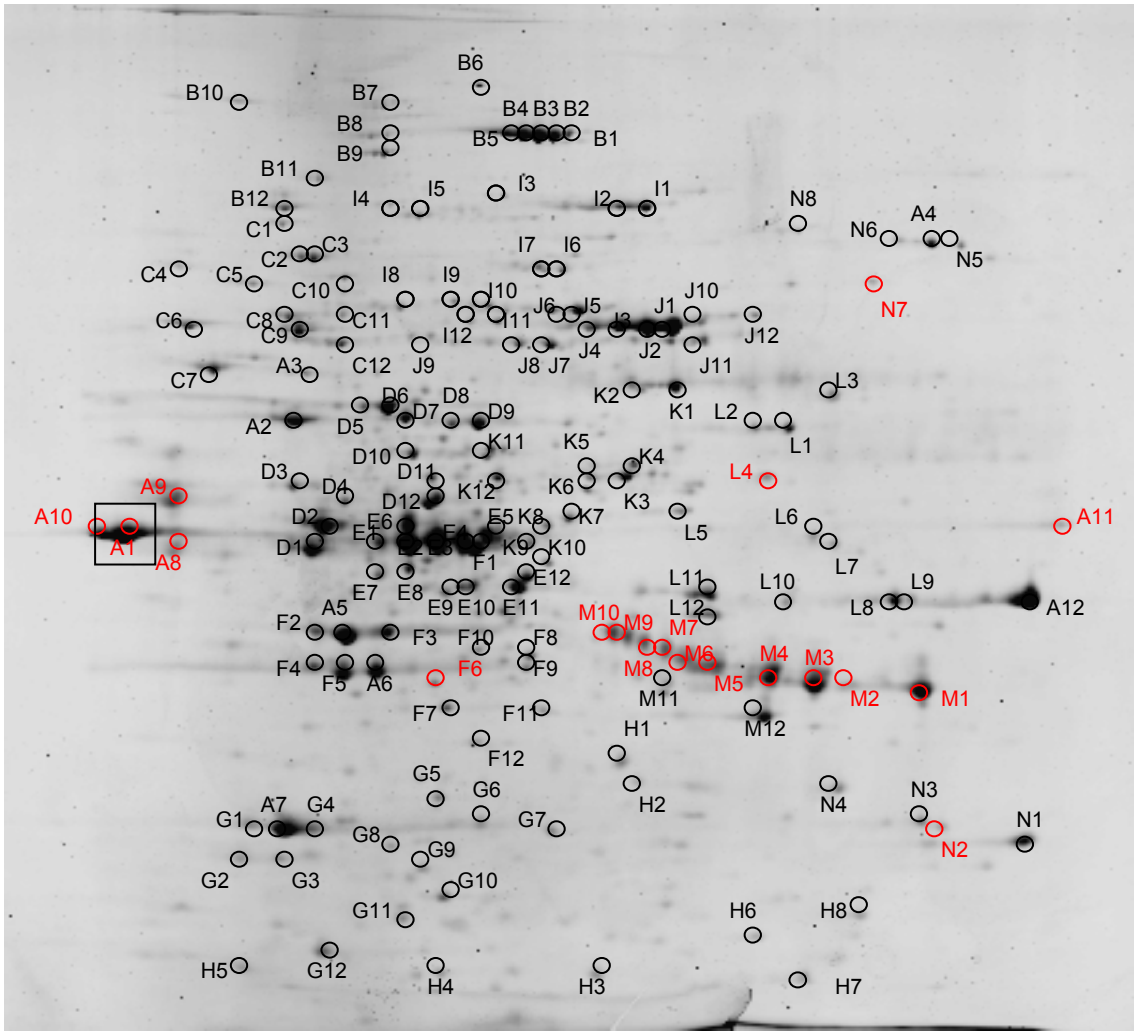


Figure 3.9 *– Spots picked for in-gel digestion and identification on the strain B-1 reference map. Identifications are shown in Appendix A1. Identity prefixes were assigned as in Figure 3.8. Surface and virulence proteins are highlighted in red, and the S-layer proteins corresponding to the SlpA gene product of the 630 reference strain are highlighted by boxes.*

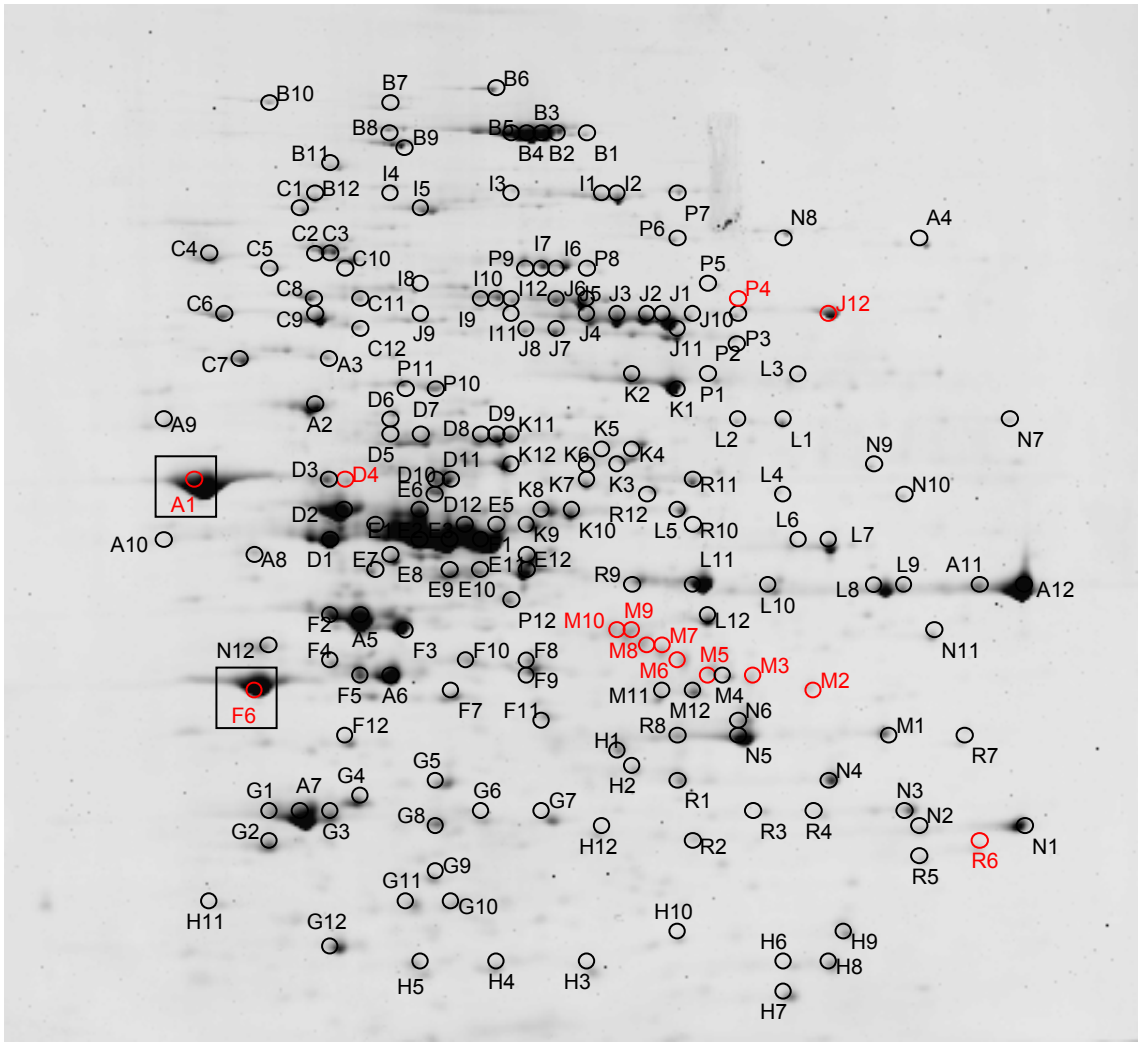


Figure 3.10 *– Spots picked for in-gel digestion and identification on the strain Tra5/5 reference map. Identifications are shown in Appendix A1. Identity prefixes were assigned as in Figure 3.8. Surface and virulence proteins are highlighted in red, and the S-layer proteins corresponding to the SlpA gene product of the 630 reference strain are highlighted by boxes.*

A breakdown of the types of proteins identified in each strain is shown in *Table 3.1*.

***Table 3.1** – Major types of proteins identified in each strain by both 1D-LC-MS/MS analysis and 2D MALDI-TOF MS analysis.*

	Hypothetical proteins		Functional proteins										Total proteins	
			Surface Virulence proteins		Metabolic proteins		Regulatory proteins		Transporter proteins		Ribosomal proteins			
	1D	**2D**	**1D**	**2D**	**1D**	**2D**	**1D**	**2D**	**1D**	**2D**	**1D**	**2D**	**1D**	**2D**
Strain A	79	6	48	7	358	57	73	6	31	1	41	2	629	80
Strain B-1	107	0	45	4	361	74	81	9	27	3	43	2	664	92
Strain Tra 5/5	36	12	23	4	271	81	49	11	17	1	29	3	425	112

A total of 154 different proteins were identified using 2D GE separation followed by MALDI-TOF MS. Only 29% of these were detected in all strains, the majority of which were metabolic enzymes. In addition, three regulatory proteins, one ribosomal proteins and the flagellin protein (FliC) were identified in all strains. A large proportion (44%) of identified proteins, were detected in only one of the three strains, including a stage IV sporulation protein A, and a putative flagella hook associated protein found in strain A only, an S-layer protein and the flagella hook protein found in strain B-1 only, and a surface protein found in strain Tra5/5 only. The majority of hypothetical proteins identified were not common to all strains.

Additional proteins of interest identified include cell surface proteins (putative S-layer proteins) and a sporulation protein Spo0A found in strains A and Tra5/5 but not B-1, and a surface protein found in B-1 and A but not Tra5/5.

3.2.2 – Protein Identification II – LC-MS/MS

LC-MS/MS has greater sensitivity than MALDI-TOF MS so was used alongside 2D reference mapping in order to increase the coverage of the proteome. The standard protein extracts were separated by 1D gel electrophoresis and each gel

lane cut into 12 bands (*Figure 3.11*), digested with trypsin and sized by LC-MS/MS. Protein identifications were obtained using MASCOT, and the results of the three strains compared using the Scaffold software (Proteome Software). Identifications are shown, along with the number of matched peptides, in *Appendix A2* and an outline of the differences between strains shown in *Figure 3.12.*

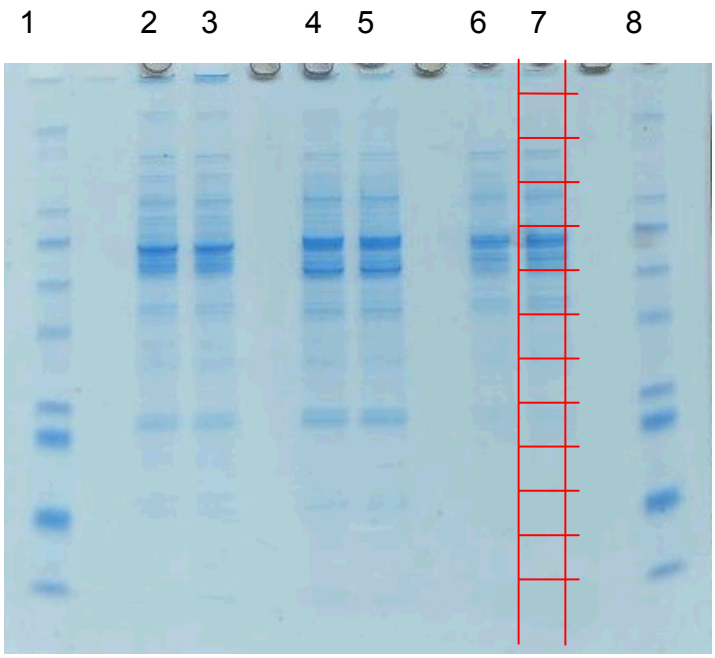
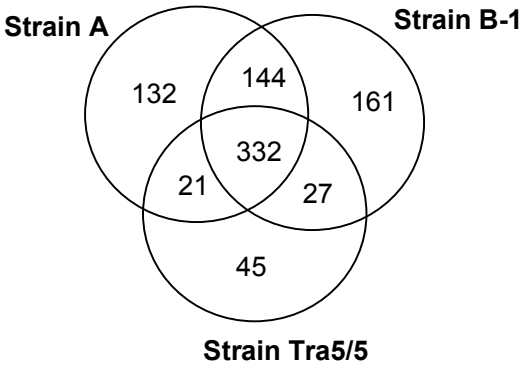


Figure 3.11 *– Coomassie stained 1D gel used for LC-MS/MS analysis. Lanes 1 and 8 are the molecular marker, lanes 2 and 3 are biological replicates of strain B-1 extract, lanes 4 and 5 are biological replicates of strain A extract and lanes 6 and 7 are biological replicates of strain Tra5/5 extract. Each lane was cut into 12 bands, as shown for lane 7, for digestion and identification.*



Total = 862 identified proteins

Figure 3.12 *– A Venn diagram showing the distribution of proteins identified by LC-MS analysis according to the strains in which they were identified.*

The 1D Gel electrophoresis LC-MS/MS workflow identified a total of 862 proteins, 734 (85%) of which had not been identified from the 2D reference maps. Many of the additional proteins identified by 1D LC-MS/MS had predicted pI values outside the range of the 2D gels, and the predicted masses of the proteins had a larger range from 4 kDa to 308 kDa.

For strain A, 629 proteins were identified, only 66 of which were seen on the strain A reference map. For strain B-1, 664 proteins were identified, 80 of which were detected on the 2D reference map, and for strain Tra5/5, fewer proteins were identified by LC-MS/MS (425). Ninety of these were confirmed by detection on 2D reference mapping.

A breakdown of the types of proteins identified in each strain is shown in *Table 3.1.*

Fewer proteins were identified by LC-MS/MS in strain Tra5/5 than in the other two strains for each general category of protein (*Table 3.1*). The same trend was not seen in the MALDI-TOF MS data, where more proteins were identified in strain Tra5/5 than in the other two strains. Both analytical methods identified a higher proportion of surface or virulence proteins in strain A than in the other two strains. By LC MS/MS analysis, 4.8% of the proteins detected in strain A were surface or virulence proteins compared to 4.1% and 4.0% of the proteins identified in strains B-1 and Tra5/5, and by MALDI-TOF MS analysis, this was (7.5%) compared with (5.3%) and (4.3 %) for strains A, B-1 and Tra5/5 respectively.

Surface and virulence proteins are likely to play an important role in determining the degree of pathogenicity of a strain, and so these proteins were examined in detail and are shown in red in *Figures 3.8, 3.9 and 3.10*. The 2D profiles of the surface and virulence proteins for the three strains are fairly similar, and comprise a number of the most abundant proteins.

No toxins were identified on the 2D reference maps which may be because large proteins are poorly resolved by 2D gel electrophoresis. Problems with visualising toxins by 2D GE have been reported previously (Mukherjee *et al*, 2002). LC-MS/MS analysis detected toxin A in all three strains, but toxin B was detected in strains A and B-1 only. Flagellin (FliC) was identified in all three strains by both MALDI-TOF MS and LC-MS/MS. On each 2D map, the flagellin protein was identified as a series of spots in the bottom right-hand quarter of the gel (the red spots with the M preface, *Figures 3.8, 3.9 and 3.10*), indicating post-translational modifications. The flagellar hook protein was identified in

strain B-1 and strain A, but was not detected at all in strain Tra5/5. A number of flagellar proteins including motor switch proteins, an M-ring protein and two flagellar biosynthesis proteins were detected in strain B-1, but neither of the other two strains, and a flagellar export protein was found only in strain A. In addition, a putative methyl accepting chemotaxis protein, gi|115249554, was detected in strains A and B-1 but not Tra5/5, and an additional chemotaxis protein, gi|115249265, was seen only in strain B-1. A number of sporulation proteins were identified in each strain, with spo0A and spoIIIJ associated protein identified in all three strains. Additional sporulation proteins were detected in strains A and B-1 but not Tra5/5.

Twenty surface proteins were identified. Stand alone BLAST was used to match the different proteins between strains, and each S-layer protein is shown in *Table 3.2* with percentage homology scores and identification numbers for each strain, and the corresponding identification number for the equivalent protein from strain 630, used as a reference. Due to the large sequence variation of surface proteins, particularly S-layer proteins, some detected proteins, such as SLP16, do not match to the 630 genome with high homology. Some of these matched to more than one different protein in the 630 reference strain with low sequence homology, or vice versa. In this case, homology is shown to more than one protein.

Table 3.2 *– Blast results of all the surface proteins identified. The percentage identification score to protein with highest homology is shown for each strain, and the 630 reference strain, along with the strains in which they were detected (I-strains in which the protein were identified by LC-MS/MS; (II)– strains in which the protein was identified by MALDI-TOF). The protein highlighted in red shows the highest sequence homology to the S-layer protein of the 630 reference strain.*

S-layer protein		630 protein ID (Sebaihia *et al*, 2006)	Strain B-1	% ID	Strain A	% ID	Strain Tra5/5	% ID	Strain I (II)
1	Cell surface protein (putative penicillin-binding protein)	gi|115250510	BASYS01592	99.9	BASYS01784	98.9	BASYS01530	99.0	A, B-1, Tra5/5 -
2	cell surface protein	gi|115251820	BASYS02999	99.9	BASYS03242	99.7	BASYS02869	99.9	A, B-1, Tra5/5 -
3	cell surface protein	gi|115251837	BASYS03016	99.6	BASYS03260	99.7	BASYS02886	99.8	A, B-1, Tra5/5 -
4	cell surface protein (putative cell surface-associated cysteine protease)	gi|115251840 (Cwp84)	BASYS03019	99.4	BASYS03263	99.6	BASYS02892	99.6	A, B-1, Tra5/5 (B-1)
5	cell surface protein	gi|115250068	BASYS01105	99.7	BASYS01335	100	BASYS01118	99.7	A, B-1, Tra5/5 -
6	putative surface protein	gi|115251769	BASYS02946	99.2	BASYS03190	99.4	BASYS02816	99.5	A, B-1, Tra5/5 -
7	putative surface protein	gi|115251764	BASYS03185	99.4	BASYS02941	99.8	BASYS02811	99.8	A, B-1 -
8	cell surface protein	gi|115249861	BASYS00545	100	BASYS01194	100	BASYS00987	100	A, B-1 -
9	Cell surface protein	gi|115251842 (Cwp66)	BASYS03070	58.9	BASYS03265	79.5	BASYS02894	79.2	A, B-1 (Tra5/5)
10	cell surface protein	gi|115251839	BASYS03018	98.4	BASYS03262	98.8	BASYS02891	99.0	A, B-1 -
11	cell wall hydrolase	gi|115251237	BASYS01205 BASYS02384	58.1 100	BASYS02646	100	BASYS02247	100	A, B-1 -
12	Cell surface protein	gi115251852	BASYS03079	99.9	BASYS03269 BASYS03276	37.0 99.5	BASYS02898 BASYS02904	36.8 99.4	A, Tra5/5 (A, Tra5/5)
13	cell surface protein (putative S-layer protein precursor	gi|115251844	BASYS03079	36.5	BASYS03267	97.9	BASYS02896	99.5	A, Tra5/5 (A, Tra5/5)

S-layer protein		630 protein ID (Sebaihia *et al*, 2006)	Strain B-1	% ID	Strain A	% ID	Strain Tra5/5	% ID	Strain I (II)
14	Cell surface protein (putative hemagglutinin/adhesin	gi|115249532	BASYS00903	100	BASYS00821	99.8	BASYS00630	100	B-1, Tra5/5 -
15	cell surface protein	gi|115251246	BASYS02392	100	BASYS02655	100	BASYS02258	100	Tra5/5 -
16	Cell surface protein (S-layer precursor protein)	gi|115251846 (SlpA)	BASYS03072	52.9	BASYS03269	55.4	BASYS02898	55.7	A, B-1, Tra5/5 (A, B-1, Tra5/5)
17	cell surface protein	gi|115251849	BASYS03075	98.3	BASYS03273	97.8	BASYS02901	99.0	B-1 -
18	putative cell wall hydrolase	gi|115252630	BASYS03924	100	BASYS00176	99.8	BASYS03714	99.4	A -
19	BASYS03271, 3480933-3479344 (CounterClockwise) Hypothetical Protein BASYS03271 Surface surface protein	gi|115251847	BASYS03073	65.8	BASYS03271	95.7	BASYS02899	95.7	A -
20	cell surface protein	gi|115251848	BASYS03074	99.1	BASYS03272	99.1	BASYS02900	99.1	A -

Eight of these surface proteins, SLP1, SLP2, SLP3, SLP4, SLP5, SLP6, SLP9 and SLP16 were detected in all three strains. Six of these showed highly conserved sequences between strains (>95% sequence homology), however SLP9 and SLP16 showed great sequence variability. Four additional proteins SLP7, SLP8, SLP10 and SLP11, were identified in strains A and B-1, but not Tra5/5. All of these showed highly conserved sequences between strains. Two proteins were identified in strain A and Tra5/5 but not B-1. SLP12 which showed high sequence homology between strains, and SLP13 which showed a highly conserved sequence between strains A and Tra5/5, but only matched with low homology to strain B-1. This match was to the same protein (BASYS03079) as SLP12, indicating that the corresponding gene to this S-layer protein may not be present in the genome of strain B-1.

One protein was detected in strains Tra5/5 and B-1 but not A, and showed a highly conserved sequence between the strains. Three proteins were detected in strain A only, one in strain B-1 only and one in strain Tra5/5 only, all of which had highly conserved sequences between strains.
The protein showing the highest homology to the 630 SlpA (CD2793, gi|115251846) is numbered 16 in *Table 3.2* above and highlighted in red. This was identified in all three strains by LC-MS/MS, and on all three reference maps as spots A1 in strain B-1, and spots A1 and F6 in strains A and Tra5/5 (boxed spots, *Figures 3.8, 3.9 and 3.10*). The protein numbered 13 in *Table 2* above (gi|115251844) was also identified in the reference maps for strains A and Tra5/5 as the spot J12 (*Figures 3.9 and 3.10*), but was not present in strain B-1. This protein was also not detected in strain B-1 by LC-MS/MS, but was in the other two strains, and as mentioned above, the corresponding gene may not be presenting the B-1 genome.

The 19 membrane proteins identified were also investigated further and are shown in *Table 3.3* below. None of these were detected on the 2D reference maps. All membrane proteins showed high homology across the three strains. The highest numbers of membrane proteins were detected in strain A, and the fewest membrane proteins were detected in strain Tra5/5. Seven proteins were detected in all three strains, five were found in strains A and B-1 only, a further six were detected only in strain A, and one was detected only in strain B-1.

Table 3.3 *– Blast results of all the membrane proteins identified. The percentage identification score to protein with highest homology is shown for each strain, and the 630 reference strain, along with the strains in which they were detected.*

Membrane Protein		630 protein ID (Sebaihia *et al*, 2006)	Strain B-1	% ID	Strain A	% ID	Strain Tra5/5	% ID	
1	putative membrane protein	gi|115249895	BASYS01231	100	BASYS00503	100	BASYS01023	100	A, B-1, Tra5/5
2	putative membrane protein	gi|115249537	BASYS00828	99.7	BASYS00886	100	BASYS00635	100	A, B-1, Tra5/5
3	putative membrane protein	gi|115250814	BASYS02178	100	BASYS01975	100	BASYS01851	100	A, B-1, Tra5/5
4	putative membrane protein	gi|115249847	BASYS01180	100	BASYS00559	100	BASYS00973	100	A, B-1, Tra5/5
5	putative membrane protein	gi|115249032	BASYS00357	99.7	BASYS00019	99.7	BASYS00024	99.7	A, B-1, Tra5/5
6	putative membrane protein	gi|115251110	BASYS02508	100	BASYS02247	99.8	BASYS02109	100	A, B-1, Tra5/5
7	putative membrane protein	gi|115250415	BASYS01679	100	BASYS01493	100	BASYS01426	100	A, B-1, Tra5/5
8	putative membrane protein	gi|115252518	BASYS00063	100	BASYS03812	100	BASYS03601	100	A, B-1
9	putative membrane protein	gi|115249607	BASYS00916	98.5	BASYS00811	100	BASYS00713	100	A, B-1
10	conserved hypothetical membrane protein	gi|115250847	BASYS02213	100	BASYS02008	100	BASYS01885	100	A, B-1
11	putative membrane protein	gi|115250318	BASYS01579	99.3	BASYS01394	100	BASYS01331	99.9	A, B-1
12	putative membrane protein	gi|115250109	BASYS01377	99.2	BASYS01225	99.7	BASYS01159	99.7	A, B-1
13	putative membrane protein	gi|115249794	BASYS01126	100	BASYS00613	100	BASYS01972	99.8	B-1
14	putative membrane protein	gi|115250974	BASYS02381	99.8	BASYS02095	100	BASYS00920	100	A
15	putative membrane protein	gi|115250482	BASYS01756	100	BASYS01564	100	BASYS01501	100	A
16	putative membrane protein	gi|115252127	BASYS03514	98.4	BASYS03296	98.9	BASYS03123	98.9	A
17	putative membrane protein	gi|115251282	BASYS02685	100	BASYS02429	100	BASYS02296	100	A
18	putative membrane protein	gi|115252285	BASYS03725	99.1	BASYS03582	99.7	BASYS03320	99.7	A
19	putative multidrug efflux pump, membrane protein	gi|115251459	BASYS02866	99.9	BASYS02621	100	BASYS02488	99.8	A

Although the 1D gel electrophoresis LC-MS methodology is more sensitive and led to the identification of a greater number of proteins, there were still 26 proteins identified in the 2D reference maps which were not identified at all by the LC-MS methodology, listed in *Appendix A3*, along with identification numbers, masses and a MASCOT identification score for each strain. These

include an indolepyruvate oxidoreductase subunit and a formate-tetrahyrdrofolate ligase which were present as abundant spots on all three 2D reference maps. Other proteins identified on the reference maps but not detected by LC-MS/MS were a number of other metabolic enzymes, a flagellar hook associated protein, and two transcription regulatory factors.

Those proteins identified as present in one strain only by LC-MS/MS were then analysed in more detail.

1D GE LC-MS/MS analysis identified 132 proteins which were present only in strain A (*Table 3.4*), 30 of which (23%) were hypothetical proteins with no putative function. Fifteen of the proteins with putative function were surface proteins or potential virulence factors including six putative membrane proteins, five sporulation related proteins, two surface proteins, one putative cell wall hydrolase and a flagella export protein. In addition, 65 of the functional proteins were metabolic enzymes, ten were transporter related proteins, 11 were regulatory proteins and one ribosomal protein was detected in strain A only.

1D GE LC-MS/MS analysis identified 161 proteins present only in strain B-1, 56 of which (35%) were hypothetical proteins with no putative function (*Table 3.5*). Of the functional proteins detected only in strain B-1, 12 were surface proteins or potential virulence factors including two cell surface proteins, one putative membrane protein, three sporulation proteins and six flagellar related proteins. In addition, 68 metabolic enzymes, five transported proteins, 17 regulatory proteins and three ribosomal proteins were detected in strain B-1 only.

1D GE LC-MS/MS analysis identified 45 proteins present in only strain Tra5/5, only six of which (13%) were hypothetical with no putative function (*Table 3.6)*. Proteins present in the Tra5/5 strain only include a cell surface protein, 33 metabolic enzymes, three regulatory proteins and two transporter proteins.

To determine whether these differences in protein expression between the strains were due to genotypic or phenotypic variation, the sequences of proteins detected in one strain only were searched against the predicted gene products of the unfinished genomes of each strain using stand alone BLAST, to determine whether the corresponding genes were present in the genome. The percentage identification and e-value scores for each protein against each genome are shown in *Tables 3.4, 3.5 and 3.6*.

The majority of the proteins identified in only one strain matched to gene products from all three genomes with high homology i.e. a percentage id greater that 90% (bold scores *Tables 3.4, 3.5 and 3.6)*. Those that did not, were mainly S-layer proteins or hypothetical proteins.

Of the 132 proteins identified only in strain A, 121 (92%) matched to gene products from all three genomes with a high homology (*Table 3.4*). Those that didn't include a surface protein (sta_BASYS03271) which matched to strain Tra5/5 with high homology, but lower homology (65%) to strain B-1. Four proteins (two hypothetical proteins, an acytransferase and a thymidylate synthase) showed homology to strain A but none to strains B-1 and Tra5/5. One conserved hypothetical protein showed 100% homology to strains A and B-1 but

much lower (32%) homology to strain Tra5/5. The others, two hypothetical proteins, a putative lipoprotein, a putative glycoxylase and a transcriptional regulator showed high homology to strain A, but lower homology to the other two strains.

Of the 161 proteins identified only in strain B-1, 135 (84%) matched to gene products from all three genomes with high homology (*Table 3.5*). Of those that didn't, four were from the strain B-1 phages, so did not match any of the strains with high homology. The S-layer precursor protein stb_BASYS03072 matched to strain B-1 with high homology, but much lower homology to strains A and Tra 5/5. Ten hypothetical proteins match to strain B-1 with high homology, but show no match to either of the other two strains. Two putative phage proteins and a hypothetical protein matched to both strains A and B-1 with high homology, but much lower homology to strain Tra5/5. The remaining six proteins showed high homology to strain B-1 but lower homology to A and Tra5/5.

Of the 45 proteins identified only in strain Tra5/5, 43 (96%) matched to gene products from all three genomes with high homology (*Table 3.6*). The two that don't are hypothetical proteins, and match to strain Tra5/5 with high homology, but much lower homology (25-30%) to strains A and B-1.

Table 3.4 *– Proteins identified in strain A only showing BLAST percentage identification and e-value scores to each of the three strains.*

	Strain A only		Strain A		Strain B-1		Strain Tra5/5	
			% ID	E value	% ID	E value	% ID	E value
293	putative bifunctional protein: peroxiredoxin/chitinase	630_gi|115250474	**100**	**0**	**99.86**	**0**	**99.86**	**0**
294	putative membrane protein	630_gi|115250974	**99.78**	**0**	**100**	**0**	**99.78**	**0**
306	probable proton-dependent oligopeptide transporter	630_gi|115251316	**100**	**0**	**100**	**0**	**100**	**0**
307	putative cell wall hydrolase	630_gi|115252630	**99.81**	**0**	**100**	**0**	**99.42**	**0**
319	single-stranded DNA binding protein	630_gi|115252292	**100**	**1.00E-119**	**100**	**1.00E-119**	**100**	**1.00E-119**
345	BASYS01831, 1920387-1919473 (CounterClockwise) Hypothetical Protein CAC Hypothetical protein	sta_BASYS01831 (115250552)	**100**	**2.00E-178**	**99.34**	**3.00E-177**	**99.67**	**8.00E-178**
382	BASYS00914, 944080-949320 (Clockwise) Hypothetical Protein Pc Hypothetical protein	sta_BASYS00914 (115249605)	**100**	**0**	**97.71**	**0**	**97.48**	**0**
424	BASYS03447, 3680327-3679458 (CounterClockwise) Hypothetical Protein CPE Hypothetical protein	sta_BASYS03447 (115252065)	**100**	**2.00E-169**	**98.27**	**2.00E-167**	**98.96**	**5.00E-168**
428	BASYS01825, 1915638-1914313 (CounterClockwise) Hypothetical Protein BASYS01825 putative drug/sodium antiporters	sta_BASYS01825 (115250546)	**100**	**0**	**99.55**	**0**	**99.32**	**0**
455	glycine betaine/carnitine/choline ABC transporter, substrate-binding protein	630_gi|115249915	**99.42**	**0**	**99.61**	**0**	**99.61**	**0**
488	putative pyruvate formate-lyase 3 activating enzyme	630_gi|115252339	**99.67**	**4.00E-177**	**100**	**9.00E-178**	**100**	**8.00E-178**
489	conserved hypothetical protein	630_gi|115252583	**94.17**	**3.00E-109**	**93.33**	**4.00E-109**	**92.92**	**6.00E-109**
490	putative oxidoreductase	630_gi|115249453	**99.15**	**0**	**99.72**	**0**	**100**	**0**
491	putative esterase/halogenase	630_gi|115251917	**99.63**	**8.00E-160**	**99.63**	**2.00E-159**	**100**	**2.00E-160**
492	BASYS03271, 3480933-3479344 (CounterClockwise) Hypothetical Protein BASYS03271 Surface surface protein	sta_BASYS03271 (115251847)	100	0	65.23	0	98.87	0
493	BASYS00861, 893869-892595 (CounterClockwise) Spore Cortex-Lytic Pre-Pro-Form	sta_BASYS00861 115249567	**100**	**0**	**99.53**	**0**	**99.53**	**0**
496	stage V sporulation protein T	630_gi|115252560	**100**	**1.00E-98**	**100**	**1.00E-98**	**100**	**1.00E-98**
497	putative protein translocase	630_gi|115251854	**100**	**5.00E-**	**100**	**5.00E-**	**100**	**5.00E-**

	subunit			**51**		**51**		**51**
537	BASYS01931, 1999997-1998948 (CounterClockwise) Hypothetical Protein BB ABC transporter substrate-binding protein [*Clostridium difficile* QCD-63q42]	stt_BASYS01931	**99.14**	**0**	**99.43**	**0**	**100**	**0**
549	glnS, 2609350-2611014 (Clockwise) Glutaminyl-tRNA synthetase	sta_BASYS02511 115251113	**100**	**0**	**99.82**	**0**	**99.82**	**0**
550	rubrerythrin	630_gi|115251898	**100**	**3.00E-115**	**100**	**3.00E-115**	**100**	**3.00E-115**
551	conserved hypothetical protein	630_gi|115252644	**99.6**	**2.00E-142**	**100**	**6.00E-143**	**99.6**	**4.00E-142**
552	phosphoribosylaminoimidazole -succinocarboxamide synthase	630_gi|115249227	**100**	**1.00E-123**	**100**	**1.00E-123**	**100**	**1.00E-123**
569	D-ornithine aminomutase E component	630_gi|115249457	**99.86**	**0**	**100**	**0**	**100**	**0**
608	putative sugar-phosphate kinase	630_gi|115249217	**98.57**	**0**	**99.76**	**0**	**99.76**	**0**
609	ABC transporter, ATP-binding protein	630_gi|115251510	**99.46**	**0**	**99.73**	**0**	**99.73**	**0**
610	thioredoxin	630_gi|115250733	**99.05**	**1.00E-57**	**99.05**	**1.00E-57**	**99.05**	**1.00E-57**
618	putative manganese-containing catalase	630_gi|115250607	**98.67**	**2.00E-131**	**99.56**	**3.00E-132**	**100**	**2.00E-132**
638	putative stage II sporulation protein P	630_gi|115251523	**99.71**	**5.00E-170**	**100**	**2.00E-170**	**100**	**2.00E-170**
639	putative aromatic amino acid aminotransferase	630_gi|115251881	**100**	**0**	**100**	**0**	**100**	**0**
640	BASYS03656, 3921940-3921434 (CounterClockwise) Acetyltransferase GNAT Family	sta_BASYS03656	100	1.00E-97	-	-	-	-
652	inosine-uridine preferring nucleoside hydrolase	630_gi|115250725	**99.69**	**0**	**99.69**	**0**	**99.38**	**0**
655	conserved hypothetical protein	630_gi|115251175	**98.82**	**1.00E-98**	**100**	**2.00E-99**	**100**	**2.00E-99**
657	deoxyuridine 5'-triphosphate nucleotidohydrolase	630_gi|115251455	**100**	**2.00E-80**	**100**	**2.00E-80**	**100**	**2.00E-80**
664	putative fumarate hydratase, subunit B	630_gi|115250025	**100**	**3.00E-92**	**100**	**3.00E-92**	**99.45**	**7.00E-92**
665	putative membrane protein	630_gi|115250482	**100**	**5.00E-147**	**100**	**5.00E-147**	**100**	**5.00E-147**
666	putative membrane protein	630_gi|115252127	**98.39**	**2.00E-79**	**98.92**	**7.00E-80**	**98.92**	**7.00E-80**
678	putative ATP/GTP-binding protein	630_gi|115252354	**99.72**	**0**	**100**	**0**	**100**	**0**
683	putative penicillin-binding protein	630_gi|115250180	**99.5**	**0**	**99.9**	**0**	**99.9**	**0**
687	PTS system, IIb component	630_gi|115250107	**98.9**	**4.00E-47**	**100**	**1.00E-47**	**100**	**1.00E-47**
688	low-specificity L-threonine aldolase	630_gi|115251648	**100**	**0**	**99.71**	**0**	**100**	**0**

689	conserved hypothetical protein	630_gi|115252296	**100**	**1.00E-72**	**100**	**1.00E-72**	**100**	**1.00E-72**
694	sodium:dicarboxylate symporter family protein	630_gi|115251595	**98.92**	**0**	**99.57**	**0**	**99.57**	**0**
695	putative lipoate-protein ligase	630_gi|115250695	**99.39**	**0**	**99.7**	**0**	**99.7**	**0**
723	putative lipoprotein	630_gi|115250265	99.27	3.00E-138	41.23	8.00E-29	43.24	3.00E-41
724	IclR-family transcriptional regulator	630_gi|115251483	**100**	**5.00E-141**	**100**	**5.00E-141**	**100**	**5.00E-141**
725	putative dihydroorotate dehydrogenase, catalytic subunit	630_gi|115252236	**100**	**0**	**100**	**0**	**100**	**0**
726	BASYS02324, 2426262-2427068 (Clockwise) Hypothetical Protein EF Hypothetical protein	sta_BASYS02324 115250916	100	2.00E-156	-	-	-	-
733	E3 component of acetoin dehydrogenase enzyme system (dihydrolipoyl dehydrogenase)	630_gi|115249042	**99.48**	**0**	**99.48**	**0**	**99.83**	**0**
734	flagellar export protein	630_gi|115249272	**99.42**	**0**	**99.42**	**0**	**99.42**	**0**
735	PTS system, IIb component	630_gi|115249506	**99.36**	**0**	**100**	**2.00E-87**	**99.36**	**2.00E-87**
736	putative transcriptional regulator	630_gi|115249644	**98.67**	**8.00E-87**	**100**	**4.00E-131**	**99.56**	**1.00E-130**
737	conserved hypothetical protein	630_gi|115250341	**100**	**1.00E-129**	**100**	**6.00E-85**	**100**	**5.00E-85**
738	putative sigma-54-dependent transcriptional regulator (partial)	630_gi|115250478	**99.49**	**1.00E-100**	**99.49**	**1.00E-100**	**99.49**	**1.00E-100**
739	putative deoxyribose-phosphate aldolase	630_gi|115250542	**100**	**5.00E-85**	**100**	**6.00E-115**	**99.54**	**4.00E-115**
740	putative ferrous ion transport protein	630_gi|115250558	**100**	**5.00E-115**	**99.86**	**0**	**100**	**0**
741	conserved hypothetical protein	630_gi|115250804	**98.91**	**0**	**100**	**5.00E-89**	**100**	**2.00E-162**
742	two-component response regulator	630_gi|115250957	**100**	**1.00E-160**	**99.48**	**5.00E-106**	**100**	**2.00E-106**
743	conserved hypothetical protein	630_gi|115251114	**100**	**2.00E-106**	**99.67**	**3.00E-170**	**100**	**2.00E-171**
744	putative membrane protein	630_gi|115251282	**100**	**2.00E-171**	**100**	**6.00E-127**	**100**	**6.00E-127**
745	putative germination-specific protease	630_gi|115251301	**99.28**	**6.00E-127**	**99.82**	**0**	**100**	**0**
746	conserved hypothetical protein	630_gi|115251313	**100**	**4.00E-61**	**100**	**4.00E-61**	**100**	**4.00E-61**
747	putative aliphatic sulfonate ABC transporter,ATP-binding protein	630_gi|115251414	**100**	**3.00E-134**	**100**	**3.00E-134**	**100**	**3.00E-134**
748	conserved hypothetical protein	630_gi|115251427	**100**	**7.00E-46**	**100**	**7.00E-46**	**100**	**7.00E-46**
749	histidine triad nucleotide-binding protein	630_gi|115251501	**100**	**3.00E-65**	**100**	**3.00E-65**	**100**	**3.00E-65**
750	30S ribosomal protein S20	630_gi|115251527	**98.86**	**4.00E-44**	**98.86**	**4.00E-44**	**98.86**	**4.00E-44**

751	putative phosphoribosyltransferase	630_gi|115251742	**100**	**4.00E-96**	**100**	**4.00E-96**	**100**	**4.00E-96**
753	dipicolinate synthase, A chain	630_gi|115252026	**99.66**	**3.00E-160**	**99.31**	**1.00E-159**	**99.66**	**3.00E-160**
754	delta-aminolevulinic acid dehydratase	630_gi|115252479	**99.69**	**0**	**100**	**0**	**100**	**0**
755	putative RNA/single-stranded DNA exonuclease	630_gi|115252723	**100**	**0**	**100**	**0**	**99.85**	**0**
756	thyA, 395829-396659 (Clockwise) Thymidylate synthase	sta_BASYS00387	100	1.00E-163	-	-	-	-
757	ligA, 3927560-3925527 (CounterClockwise) Hypothetical Protein ligA DNA ligase	stb_BASYS03670 (115252365)	**99.85**	**0**	**100**	**0**	**99.56**	**0**
758	ycjC, 1769550-1768987 (CounterClockwise) Hypothetical Protein ycjC	stt_BASYS01717 115250685	**100**	**8.00E-100**	**99.47**	**4.00E-99**	**100**	**8.00E-100**
759	kdgR, 3236380-3237156 (Clockwise) Hypothetical Protein kdgR	stt_BASYS03056 115249418	**100**	**4.00E-144**	**100**	**4.00E-144**	**100**	**4.00E-144**
763	putative exported protein	630_gi|115252517	**99.28**	**3.00E-74**	**99.28**	**3.00E-74**	**99.28**	**3.00E-74**
767	yecS, 1915939-1916550 (Clockwise) Hypothetical Protein yecS amino acid ABC transporter, permease protein	stt_BASYS01858 (115250821)	**100**	**2.00E-113**	**100**	**2.00E-113**	**100**	**2.00E-113**
770	stage V sporulation protein S	630_gi|115250981	**100**	**4.00E-44**	**100**	**4.00E-44**	**100**	**4.00E-44**
771	BASYS00233, 228947-229429 (Clockwise) Glyoxalase	sta_BASYS00233	100	1.00E-92	52.6	8.00E-43	51.95	5.00E-42
832	putative glyoxalase	630_gi|115249166						
773	PTS system, mannitol-specific IIbc component	630_gi|115251389	**99.79**	**0**	**99.58**	**0**	**99.37**	**0**
815	bifunctional purine biosynthesis protein includes: phosphoribosylaminoimidazole carboxamide formyltransferase and IMP cyclohydrolase	630_gi|115249231	**99.41**	**0**	**100**	**0**	**100**	**0**
816	methylglyoxal synthase	630_gi|115250185	**100**	**0**	**100**	**1.00E-78**	**100**	**1.00E-78**
817	putative muramoyltetrapeptide carboxypeptidase	630_gi|115250438	**99.65**	**1.00E-78**	**99.65**	**1.00E-166**	**100**	**4.00E-167**
818	cell surface protein	630_gi|115251848	**99.06**	**1.00E-166**	**99.06**	**0**	**99.06**	**0**
819	conserved hypothetical protein	630_gi|115252555	**100**	**0**	**100**	**4.00E-44**	**100**	**4.00E-44**
820	manX, 3750702-3750208 (CounterClockwise) Hypothetical Protein manX PTS system, IIB component	sta_BASYS03509 (11525212)	**100**	**4.00E-44**	**100**	**1.00E-92**	**100**	**1.00E-92**

821	BASYS03604, 3866978-3865935 (CounterClockwise) Hypothetical Protein BASYS03604 Hypothetical protein	sta_BASYS03604	100	1.00E-92	29.34	4.00E-22	29.34	3.00E-22
831	putative exported protein	630_gi|115250383	**98.03**	**0**	**93.54**	**1.00E-153**	**99.72**	**1.00E-167**
833	putative spore-coat protein	630_gi|115249613	**99.47**	**2.00E-164**	**100**	**2.00E-113**	**100**	**2.00E-113**
834	conserved hypothetical protein	630_gi|115250402	**98.65**	**9.00E-113**	**100**	**7.00E-82**	**98.65**	**1.00E-80**
835	putative signaling protein	630_gi|115250461	**99.65**	**1.00E-80**	**99.83**	**0**	**99.83**	**0**
836	putative molybdenum cofactor biosynthesis protein	630_gi|115250756	**100**	**0**	**100**	**6.00E-93**	**100**	**6.00E-93**
837	putative glutamyl-aminopeptidase	630_gi|115251205	**100**	**6.00E-93**	**100**	**0**	**100**	**0**
838	putative oxidoreductase subunit	630_gi|115251480	**100**	**0**	**100**	**2.00E-127**	**100**	**2.00E-127**
839	conserved hypothetical protein	630_gi|115251531	**100**	**2.00E-127**	**100**	**2.00E-166**	**99.67**	**1.00E-165**
840	uracil-DNA glycosylase	630_gi|115251535	**100**	**2.00E-166**	**100**	**9.00E-133**	**100**	**9.00E-133**
841	PyrR bifunctional protein includes: pyrimidine operon regulatory protein; uracil phosphoribosyltransferase	630_gi|115251652	**100**	**9.00E-133**	**100**	**4.00E-97**	**100**	**3.00E-97**
842	tryptophanyl-tRNA synthetase	630_gi|115251661	**99.7**	**3.00E-97**	**99.7**	**0**	**100**	**0**
843	putative ATP-binding protein	630_gi|115251691	**99.35**	**0**	**100**	**3.00E-75**	**100**	**3.00E-75**
844	oligopeptide ABC transporter, permease protein	630_gi|115251724	**100**	**1.00E-74**	**100**	**4.00E-170**	**99.69**	**2.00E-169**
845	probable polysaccharide deacetylase	630_gi|115251771	**99.68**	**4.00E-170**	**100**	**1.00E-171**	**100**	**1.00E-171**
846	conserved hypothetical protein	630_gi|115251862	100	3.00E-57	100	3.00E-57	32.94	5.00E-07
847	conserved hypothetical protein	630_gi|115252126	**98.27**	**8.00E-90**	**100**	**7.00E-91**	**98.27**	**8.00E-89**
848	putative flavodoxin	630_gi|115252176	**98.39**	**3.00E-138**	**99.6**	**4.00E-140**	**99.6**	**4.00E-140**
849	probable sigma54-dependent transcriptional regulator	630_gi|115252242	**99.17**	**0**	**99.83**	**0**	**99.83**	**0**
850	putative membrane protein	630_gi|115252285	**99.05**	**8.00E-100**	**99.7**	**1.00E-160**	**99.7**	**1.00E-160**
851	hypoxanthine phosphoribosyltransferase	630_gi|115252288	**99.44**	**3.00E-101**	**99.44**	**3.00E-101**	**99.44**	**3.00E-101**
852	probable GTP-binding protein	630_gi|115252356	**100**	**8.00E-115**	**100**	**9.00E-115**	**100**	**8.00E-115**
853	GntR-family transcriptional regulator	630_gi|115252628	**99.54**	**5.00E-125**	**99.54**	**5.00E-125**	**99.54**	**5.00E-125**
854	conserved hypothetical protein	630_gi|115252676	**100**	**3.00E-76**	**100**	**3.00E-76**	**100**	**3.00E-76**

855	BASYS00041, 40158-39427 (CounterClockwise) Hypothetical Protein BASYS00041 Hypothetical protein	sta_BASYS00041 (115249290)	100	3.00E-102	87.56	1.00E-84	87.56	1.00E-84
856	BASYS00235, 229993-230559 (Clockwise) Hypothetical Protein BASYS00235 signal peptide [*Clostridium difficile* QCD-66c26]	sta_BASYS00235	100	4.00E-86	-	-	-	-
860	BASYS03583, 3840546-3841103 (Clockwise) Transcriptional Regulator PadR-Like Family	sta_BASYS03583	100	2.00E-96	39.78	1.00E-11	39.78	9.00E-12
861	msbA, 3084615-3082825 (CounterClockwise) Hypothetical Protein msbA ABC transporter, ATP-binding/permease protein	stt_BASYS02922 (115251870)	**100**	**0**	**100**	**0**	**100**	**0**
869	ydgP, 1509321-1509890 (Clockwise) Electron Transport Complex Protein RnfG	sta_BASYS01412 115250170	**100**	**2.00E-92**	**96.83**	**5.00E-88**	**97.35**	**5.00E-89**
870	conserved hypothetical protein	630_gi|115249704	**99.43**	**6.00E-101**	**98.86**	**8.00E-101**	**98.86**	**7.00E-101**
872	putative transcriptional regulator	630_gi|115250151	**99.45**	**9.00E-88**	**100**	**2.00E-88**	**100**	**2.00E-88**
875	putative membrane-associated neutral zinc metallopeptidase	630_gi|115251638	**100**	**1.00E-119**	**100**	**1.00E-119**	**100**	**1.00E-119**
879	arginyl-tRNA synthetase	630_gi|115249727	**100**	**0**	**100**	**0**	**100**	**0**
881	putative multidrug efflux pump, membrane protein	630_gi|115251459	**99.9**	**0**	**100**	**0**	**99.8**	**0**
884	putative arsenate reductase	630_gi|115250823	**100**	**1.00E-64**	**100**	**1.00E-64**	**100**	**1.00E-64**
885	conserved hypothetical protein	630_gi|115249717	**99.49**	**2.00E-108**	**100**	**8.00E-109**	**100**	**7.00E-109**
894	aspartokinase	630_gi|115250358	**99.75**	**0**	**100**	**0**	**100**	**0**
897	ABC transporter, ATP-binding/permease protein	630_gi|115250108	**100**	**0**	**100**	**0**	**100**	**0**
899	putative sensor histidine kinase	630_gi|115251546	**99.78**	**0**	**99.89**	**0**	**100**	**0**
905	penicillin-binding protein	630_gi|115249798	**99.5**	**0**	**99.87**	**0**	**99.6**	**0**
907	putative gluconate permease	630_gi|115250781	**100**	**0**	**100**	**0**	**100**	**0**
916	DeoR-family transcriptional regulator (fatty acid and phospholipid biosynthesis regulator)	630_gi|115250210	**100**	**2.00E-104**	**100**	**3.00E-104**	**100**	**2.00E-104**
917	PTS system, mannitol-specific IIa component	630_gi|115251387	**100**	**2.00E-77**	**100**	**2.00E-77**	**100**	**2.00E-77**
919	BASYS03298, 3515475-3515182 (CounterClockwise) Hypothetical Protein BASYS03298 Hypothetical protein	sta_BASYS03298	**100**	**2.00E-24**	**97.3**	**8.00E-15**	**97.3**	**8.00E-15**
927	putative AMP-binding protein	630_gi|115252027	**97.62**	**0**	**98.7**	**0**	**98.49**	**0**
928	putative histidyl-tRNA synthetase	630_gi|115251792	**100**	**0**	**99.76**	**0**	**100**	**0**
929	sta_BASYS02525-R	sta_BASYS02525	**100**	**0**	**99.74**	**0**	**99.74**	**0**

Table 3.5 *– Proteins identified in strain B-1 only showing BLAST percentage identification and e-value scores to each of the three strains.*

	Strain B-1 only		Strain A		Strain B-1		Strain Tra5/5	
			% ID	E value	% ID	E value	% ID	E value
33	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072 Cell surface protein (S-layer precursor protein)	stb_BASYS03072 (115251846)	52.71	6.00E-113	100	0	52.91	2.00E-113
168	putative carbon-nitrogen hydrolase	630_gi|115251789	**99.25**	**7.00E-157**	**100**	**5.00E-158**	**99.63**	**2.00E-157**
255	transcription antiterminator	630_gi|115252086	**99.64**	**1.00E-154**	**100**	**5.00E-155**	**100**	**5.00E-155**
317	BASYS00003, 1394-2224 (Clockwise) Hypothetical Protein BASYS00003	Stbphage1_BASYS00003	-	-	-	-	-	-
343	conserved hypothetical protein	630_gi|115252214	**99.64**	**5.00E-159**	**100**	**2.00E-159**	**99.64**	**5.00E-159**
344	ybhF, 1494368-1495099 (Clockwise) Hypothetical Protein ybhF putative lantibiotic ABC transporter,ATP-binding protein	stb_BASYS01462 (115250385)	**97.94**	**6.00E-137**	**100**	**2.00E-139**	**99.18**	**1.00E-138**
361	sopA, 5108-4290 (CounterClockwise) Hypothetical Protein sopA	Stbphage1_BASYS00008	29.11	1.00E-11	29.11	1.00E-11	28.64	2.00E-11
379	ybaB, 338454-338834 (Clockwise) Hypothetical Protein ybaB conserved hypothetical protein	sta_BASYS00340 (115249020)	**100**	**1.00E-56**	**100**	**2.00E-56**	**100**	**1.00E-56**
397	putative oxidoreductase, electron transfer subunit	630_gi|115249185	**99.36**	**2.00E-88**	**100**	**7.00E-89**	**98.73**	**7.00E-88**
398	putative L-threonine dehydrogenase	630_gi|115250838	**99.37**	**0**	**100**	**0**	**99.69**	**0**
858	galE, 2294153-2293197 (CounterClockwise) Hypothetical Protein gale putative L-threonine dehydrogenase	sta_BASYS02203 (115250838)	**100**	**0**	**99.37**	**0**	**99.06**	**0**
420	BASYS02421, 2523005-2522130 (CounterClockwise) Hypothetical Protein BASYS02421 Hypothetical protein	sta_BASYS02421 (115251013)	**100**	**1.00E-166**	**99.66**	**2.00E-166**	**99.66**	**2.00E-166**

423	BASYS00597, 615300-615635 (Clockwise) Hypothetical Protein BASYS00597 Hypothetical protein	sta_BASYS00597 (115249246)	**100**	**7.00E-57**	**99.1**	**8.00E-56**	**99.1**	**8.00E-56**
425	BASYS00012, 9275-9622 (Clockwise) Hypothetical Protein BASYS00012	Stbphage1_BASYS00012						
426	putative phosphosugar isomerase	630_gi|115252083	**100**	**1.00E-114**	**100**	**1.00E-114**	**100**	**1.00E-114**
429	dihydropteroate synthase	630_gi|115250491	**99.63**	**1.00E-154**	**100**	**5.00E-155**	**99.63**	**8.00E-154**
430	conserved hypothetical protein	630_gi|115249635	**100**	**8.00E-38**	**100**	**8.00E-38**	**100**	**8.00E-38**
453	PTS system, IIabc component	630_gi|115252143	**99.58**	**0**	**100**	**0**	**100**	**0**
474	putative 3-mercaptopyruvate sulfurtransferase	630_gi|115250576	**100**	**2.00E-164**	**100**	**2.00E-164**	**99.65**	**1.00E-163**
482	BASYS01010, 1054756-1055364 (Clockwise) Hypothetical Protein BASYS01010 Conserved hypothetical protein	sta_BASYS01010 (115249688)	**100**	**5.00E-107**	**100**	**7.00E-108**	**100**	**6.00E-108**
484	BASYS00060, 38456-38932 (Clockwise) Hypothetical Protein BASYS00060 Hypothetical protein	stbphage2_BASYS00060					72.78	9.00E-48
486	thioredoxin	630_gi|115251408	**100**	**3.00E-58**	**100**	**4.00E-58**	**100**	**3.00E-58**
487	conserved hypothetical protein	630_gi|115249636	**100**	**2.00E-35**	**100**	**2.00E-35**	**100**	**2.00E-35**
494	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	630_gi|115252499	**100**	**0**	**100**	**0**	**100**	**0**
503	conserved hypothetical protein	630_gi|115251009	**98.86**	**1.00E-154**	**98.86**	**1.00E-154**	**98.86**	**1.00E-154**
504	peptide chain release factor 1	630_gi|115252544	**100**	**0**	**100**	**0**	**100**	**0**
524	BASYS02039, 2122887-2123531 (Clockwise) Hypothetical Protein BASYS02039 Tellurium resistance protein	sta_BASYS02039 (115250677)	**100**	**1.00E-112**	**100**	**1.00E-112**	**100**	**1.00E-112**
525	rRNA adenine N-6-methyltransferase (erythromycin resistance protein)	630_gi|115251058	27.51	9.00E-16	97.96	1.00E-139	27.51	9.00E-16
531	putative membrane protein	630_gi|115249794	**100**	**1.00E-120**	**100**	**1.00E-120**	**100**	**1.00E-120**
533	conserved hypothetical protein	630_gi|115252657	**99.49**	**1.00E-112**	**99.49**	**2.00E-112**	**100**	**3.00E-113**
534	BASYS00477, 430146-431189 (Clockwise) Hypothetical Protein BASYS00477 No c. difficile match by BLAST	stb_BASYS00477			100	1.00E-172		
535	thioredoxin reductase	630_gi|115250734	**100**	**5.00E-173**	**100**	**5.00E-173**	**99.66**	**7.00E-173**
536	putative ATP-binding protein	630_gi|115250457	**100**	**2.00E-129**	**99.55**	**4.00E-129**	**99.55**	**5.00E-129**

541	BASYS03056, 3218968-3215420 (CounterClockwise) Hypothetical Protein BASYS03056	stb_BASYS03056			100	0		
544	PTS system, IIa component	630_gi|115251565	**98.75**	**2.00E-87**	**100**	**2.00E-88**	**100**	**1.00E-88**
562	cobalt-precorrin-6a reductase; precorrin-6x reductase	630_gi|115252483	**100**	**2.00E-142**	**100**	**2.00E-142**	**100**	**2.00E-142**
564	signal recognition particle protein	630_gi|115250286	**99.56**	**0**	**99.78**	**0**	**99.78**	**0**
566	50S ribosomal protein L23	630_gi|115249079	**98.96**	**1.00E-51**	**100**	**5.00E-52**	**100**	**5.00E-52**
567	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	630_gi|115250983	**99.65**	**3.00E-169**	**100**	**1.00E-169**	**100**	**1.00E-169**
571	putative flagellar motor switch protein	630_gi|115249279	**98.63**	**9.00E-170**	**98.98**	**1.00E-169**	**98.98**	**1.00E-169**
702	fliN, 608925-609239 (Clockwise) Hypothetical Protein fliN flagellar motor switch protein	sta_BASYS00587 (115249236)	**100**	**2.00E-53**	**98.08**	**7.00E-53**	**98.08**	**7.00E-53**
710	putative flagellar motor switch protein	630_gi|115249280	**99.19**	**9.00E-63**	**100**	**3.00E-63**	**100**	**3.00E-63**
575	conserved hypothetical protein	630_gi|115250343	**100**	**4.00E-49**	**100**	**4.00E-49**	**100**	**4.00E-49**
576	probable esterase	630_gi|115251258	**98.85**	**1.00E-153**	**99.23**	**3.00E-154**	**99.47**	**1.00E-111**
577	conserved hypothetical protein	630_gi|115249911	**98.28**	**4.00E-60**	**100**	**5.00E-61**	**100**	**5.00E-61**
578	50S ribosomal protein L27	630_gi|115250195	**100**	**2.00E-43**	**100**	**2.00E-43**	**100**	**2.00E-43**
580	pyrroline-5-carboxylate reductase	630_gi|115250536	**99.63**	**3.00E-152**	**100**	**7.00E-153**	**100**	**7.00E-153**
581	glycine/sarcosine/betaine reductase complex component A	630_gi|115251406	**100**	**2.00E-27**	**100**	**2.00E-27**	**100**	**2.00E-27**
582	putative transcription antiterminator	630_gi|115251719	**99.29**	**2.00E-146**	**99.64**	**2.00E-146**	**99.29**	**8.00E-146**
583	cell surface protein	630_gi|115251849	**97.77**	**0**	**98.25**	**0**	**99.04**	**0**
584	PTS system, maltose and glucose-specific IIbc component	630_gi|115252085	**100**	**0**	**99.81**	**0**	**99.81**	**0**
585	putative stage 0 sporulation protein	630_gi|115252738	**100**	**3.00E-149**	**100**	**3.00E-149**	**100**	**3.00E-149**
586	adhE, 2136326-2137795 (Clockwise) Hypothetical Protein adhE putative ethanolamine/propanediol utilisation aldehyde-alcohol dehydrogenase	stb_BASYS02079 (115250964)	**99.39**	**0**	**100**	**0**	**99.39**	**0**
587	BASYS03023, 3180249-3179710 (CounterClockwise) Hypothetical Protein Lmo Conjugative transposon conserved hypothetical protein	stb_BASYS03023 (115249366)	63.57	2.00E-46	100	1.00E-104	41.3	5.00E-27

589	BASYS03016, 3189040-3188357 (CounterClockwise) Hypothetical Protein BF Conserved hypothetical protein	stt_BASYS03016 (115252021)	**99.56**	**2.00E-130**	**99.56**	**2.00E-130**	**100**	**2.00E-130**
604381	BASYS00474, 429093-429458 (Clockwise) Hypothetical Protein BASYS00474 Putative phage regulatory protein (XRE family transcriptional regulator) BASYS00015, 10793-10440 (CounterClockwise) Hypothetical Protein BASYS00015 Transcriptional regulator, xre family protein	stb_BASYS00474 (115251071)	47.89	2.00E-14	100	2.00E-56	47.89	2.00E-14
		Stbphage1_BASYS00015					43.94	5.00E-10
611	transcription elongation factor	630_gi|115252616	**100**	**3.00E-74**	**100**	**3.00E-74**	**100**	**3.00E-74**
612	30S ribosomal protein S16	630_gi|115250287	**100**	**5.00E-48**	**100**	**5.00E-48**	**100**	**5.00E-48**
613	Flavodoxin	630_gi|115251047	**99.3**	**8.00E-70**	**100**	**9.00E-71**	**99.3**	**8.00E-70**
619	rffG, 606774-607757 (Clockwise) Hypothetical Protein rffG (dTDP-glucose 4,6-dehydratase)	sta_BASYS00584	**100**	**0**	**97.86**	**0**	**97.86**	**0**
620	homoserine kinase	630_gi|115251173	**99.33**	**2.00E-173**	**100**	**1.00E-174**	**100**	**1.00E-174**
624	putative small-molecule-binding protein	630_gi|115249015	**86.96**	**1.00E-06**	**99.43**	**4.00E-87**	**100**	**3.00E-87**
625	ATP-dependent nuclease subunit A	630_gi|115250062	**99.37**	**0**	**99.84**	**0**	**99.76**	**0**
626	putative tRNA modification GTPase	630_gi|115252740	**99.33**	**0**	**100**	**0**	**100**	**0**
627	BASYS03349, 3561478-3560975 (CounterClockwise) Hypothetical Protein BASYS03349 No match by BLAST	stb_BASYS03349			100	4.00E-98		
644	cysG, 9043-8270 (CounterClockwise) Hypothetical Protein cysG precorrin-4 C(11)-methyltransferase	sta_BASYS00009 (115252486)	**100**	**1.00E-151**	**100**	**1.00E-151**	**100**	**1.00E-151**
645	BASYS00154, 142715-143050 (Clockwise) Hypothetical Protein BASYS00154 Phage protein	stb_BASYS00154 (115249958)			100	7.00E-62		
649	PTS system, IIa component	630_gi|115252072	**99.35**	**9.00E-83**	**100**	**3.00E-83**	**100**	**2.00E-83**
650	stage V sporulation protein G	630_gi|115252577	**100**	**1.00E-50**	**100**	**1.00E-50**	**100**	**1.00E-50**
651	hypothetical protein	630_gi|115252299	**100**	**4.00E-50**	**100**	**5.00E-50**	**100**	**4.00E-50**
653	hypothetical protein	630_gi|115251805	**99.55**	**7.00E-126**	**99.55**	**8.00E-126**	**99.55**	**7.00E-126**

654	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	630_gi|115250982	**100**	**1.00E-165**	**100**	**1.00E-165**	**100**	**1.00E-165**
660	putative exported protein	630_gi|115251418	**100**	**1.00E-63**	**100**	**1.00E-63**	**100**	**1.00E-63**
661	conserved hypothetical protein	630_gi|115251421	**100**	**4.00E-62**	**100**	**4.00E-62**	**100**	**4.00E-62**
662	BASYS00184, 156070-156327 (Clockwise) Hypothetical Protein BASYS00184 No match to BLAST	stb_BASYS00184			100	1.00E-47		
663	BASYS01737, 1767940-1768269 (Clockwise) Hypothetical Protein BASYS01737 No match to BLAST	stb_BASYS01737			100	6.00E-45		
667	conserved hypothetical protein	630_gi|115249350	**99.31**	**5.00E-79**	**99.31**	**5.00E-79**	**100**	**1.00E-79**
668	conserved hypothetical protein	630_gi|115249339	**99.6**	**2.00E-142**	**100**	**2.00E-142**	**100**	**2.00E-142**
669	putative chorismate biosynthesis-related protein includes: phospho-2-dehydro-3-deoxyheptonate aldolase and chorismate mutase	630_gi|115250494	**99.23**	**3.00E-137**	**99.62**	**1.00E-137**	**100**	**3.00E-138**
670	cysteine synthase A	630_gi|115250707	**100**	**0**	**100**	**0**	**99.69**	**0**
671	BASYS03069, 3232315-3231857 (CounterClockwise) Hypothetical Protein BASYS03069 No match by BLAST	stb_BASYS03069			100	2.00E-83		
682	tellurium resistance protein	630_gi|115250846	**100**	**1.00E-107**	**100**	**1.00E-107**	**100**	**1.00E-107**
692	putative serine hydroxymethyltransferase	630_gi|115251778	**99.28**	**0**	**99.76**	**0**	**99.76**	**0**
693	conserved hypothetical protein	630_gi|115250202	**100**	**0**	**100**	**0**	**100**	**0**
696	putative phage positive regulator of late transcription	630_gi|115249641	**100**	**5.00E-39**	**100**	**6.00E-39**	**100**	**5.00E-39**
697	conserved hypothetical protein	630_gi|115250817	**100**	**7.00E-56**	**100**	**7.00E-56**	**100**	**7.00E-56**
698	PTS system, glucose-specific IIa component	630_gi|115252082	**98.54**	**9.00E-74**	**99.27**	**1.00E-74**	**100**	**6.00E-75**
699	putative 5-formyltetrahydrofolate cyclo-ligase	630_gi|115252516	**99.47**	**4.00E-97**	**98.94**	**2.00E-97**	**100**	**1.00E-97**
700	ATP synthase epsilon chain (partial)	630_gi|115252527	**100**	**4.00E-45**	**100**	**4.00E-45**	**100**	**4.00E-45**
703	BASYS00188, 157642-158154 (Clockwise) Hypothetical Protein BASYS00188 Putative phage repressor	stb_BASYS00188	37.78	1.00E-07	100	1.00E-79	37.78	1.00E-07
704	ydhO, 1281410-1282705 (Clockwise) Hypothetical Protein ydhO putative cell wall hydrolase	stb_BASYS01250 (115250166)	**99.07**	**0**	**100**	**0**	**99.3**	**0**

	rfbC, 3226199-3225627 (CounterClockwise) Hypothetical Protein rfbC (dTDP-4-dehydrorhamnose 3,5-epimerase)	stb_BASYS03063	65.66	1.00E-64	100	4.00E-99	65.06	4.00E-64
705891	rfbC, 606179-606736 (Clockwise) Hypothetical Protein rfbC dTDP-4-dehydrorhamnose 3,5-epimerase	sta_BASYS00583						
706	BASYS03448, 3677647-3677129 (CounterClockwise) Hypothetical Protein BASYS03448 Putative nitroreductase	stt_BASYS03448 (115252412)	**98.84**	**2.00E-92**	**100**	**4.00E-93**	**100**	**4.00E-93**
708	DNA polymerase III subunit gamma/tau	630_gi|115249019	**99.82**	**0**	**100**	**0**	**100**	**0**
709	dihydroorotate dehydrogenase electron transfer subunit	630_gi|115249195	**100**	**3.00E-133**	**99.84**	**0**	**100**	**3.00E-133**
711	putative nucleotide phosphodiesterase	630_gi|115249706	**99.37**	**0**	**99.83**	**0**	**99.68**	**0**
712	conserved hypothetical protein	630_gi|115250366	**99.83**	**0**	**100**	**1.00E-57**	**99.83**	**0**
713	putative dinitrogenase iron-molybdenum cofactor	630_gi|115250736	**100**	**1.00E-57**	**100**	**2.00E-83**	**100**	**1.00E-57**
714	ABC transporter, ATP-binding protein	630_gi|115251007	**99.1**	**5.00E-124**	**99.55**	**2.00E-124**	**99.55**	**1.00E-124**
715	putative aromatic compounds hydrolase	630_gi|115251234	**99.61**	**9.00E-149**	**99.21**	**3.00E-148**	**99.61**	**9.00E-149**
717	probable cobalt-precorrin-6y C(5)-methyltransferase	630_gi|115252488	**98.51**	**7.00E-110**	**100**	**6.00E-111**	**100**	**6.00E-111**
718	carbamoyl-phosphate synthase,pyrimidine-specific, small chain	630_gi|115252653	**100**	**0**	**100**	**0**	**100**	**0**
719	sporulation initiation inhibitor	630_gi|115252737	**99.59**	**7.00E-140**	**99.59**	**7.00E-140**	**100**	**3.00E-140**
720	BASYS00599, 616624-618843 (Clockwise) Glycosyl Transferase Group 2 Family Protein	sta_BASYS00599 (115249248)	**100**	**0**	**96.75**	**0**	**96.75**	**0**
721	BASYS00167, 150151-149381 (CounterClockwise) Hypothetical Protein BASYS00167 hypothetical	stb_BASYS00167			100	2.00E-146		
722	BASYS01138, 1196854-1196450 (CounterClockwise) Transcriptional Regulator Cro/CI Family	stb_BASYS01138	68.75	2.00E-20	100	2.00E-72	35.4	1.00E-08
761	MarR-family transcriptional regulator	630_gi|115250333	**100**	**1.00E-78**	**100**	**1.00E-78**	**100**	**1.00E-78**
766	rpoN, 3932428-3931073 (CounterClockwise) Hypothetical Protein rpoN RNA polymerase sigme-54 factor	sta_BASYS03664 (115252233)	**100**	**0**	**100**	**0**	**100**	**0**

775	conserved hypothetical protein	630_gi|115250436	**99.37**	**2.00E-88**	**99.37**	**2.00E-88**	**100**	**7.00E-89**
776	hypothetical phage protein	630_gi|115249959			100	1.00E-25		
777	putative phage major capsid protein	630_gi|115249966	98.31	2.00E-162	91.97	4.00E-156	23.05	9.00E-09
778	ABC transporter, substrate-binding lipoprotein	630_gi|115250020	**99.09**	**8.00E-171**	**99.7**	**3.00E-171**	**99.7**	**3.00E-171**
779	3-oxoacyl-acyl-carrier-protein synthase III	630_gi|115250212	**100**	**0**	**100**	**0**	**100**	**0**
780	hypothetical protein	630_gi|115250621	**94.16**	**2.00E-39**	**100**	**3.00E-43**	**100**	**2.00E-43**
781	conserved hypothetical protein	630_gi|115251130	**95.32**	**0**	**87.05**	**3.00E-164**	**92.29**	**2.00E-179**
782	conserved hypothetical protein	630_gi|115251487	**100**	**3.00E-92**	**100**	**4.00E-92**	**100**	**3.00E-92**
783	putative exported protein	630_gi|115251537	**99.43**	**0**	**99.43**	**0**	**100**	**0**
784	coenzyme A biosynthesis bifunctional protein	630_gi|115251643	**99.5**	**0**	**99.75**	**0**	**99.75**	**0**
785	conserved hypothetical protein	630_gi|115252035	**98.26**	**1.00E-67**	**99.42**	**1.00E-67**	**99.22**	**2.00E-51**
786	putative preprotein translocase	630_gi|115252635	**100**	**3.00E-85**	**99.34**	**9.00E-85**	**100**	**3.00E-85**
787	conserved hypothetical protein	630_gi|115252672	**100**	**4.00E-180**	**99.67**	**2.00E-179**	**100**	**4.00E-180**
788	gidB, 310295-309567 (CounterClockwise) Hypothetical Protein gidB methyl transferase	sta_BASYS00316	**100**	**5.00E-62**	**100**	**5.00E-62**	**100**	**5.00E-62**
789	fliF, 335693-334143 (CounterClockwise) Hypothetical Protein fliF flagellar M-ring protein	stb_BASYS00356 (115249256)	**98.26**	**0**	**100**	**0**	**100**	**0**
794	trehalose-6-phosphate hydrolase	630_gi|115252145	**99.64**	**0**	**100**	**0**	**99.64**	**0**
796	chemotaxis protein	630_gi|115249265	**93.75**	**2.00E-134**	**100**	**1.00E-142**	**100**	**1.00E-142**
797	conserved hypothetical protein	630_gi|115249689	**100**	**3.00E-138**	**100**	**3.00E-138**	**100**	**3.00E-138**
798	hypothetical protein	630_gi|115249753	**100**	**1.00E-31**	**100**	**2.00E-31**	**100**	**1.00E-31**
799	putative ribonucleotide-diphosphate reductase	630_gi|115250295	**99.73**	**0**	**100**	**0**	**100**	**0**
800	630_gi|115250446|emb|CAJ68269.1|-R	630_gi|115250446	**99.25**	**1.00E-136**	**98.88**	**6.00E-136**	**99.63**	**7.00E-137**
801	molybdenum cofactor biosynthesis protein	630_gi|115250758	**99.36**	**7.00E-87**	**100**	**1.00E-87**	**100**	**1.00E-87**
802	DNA mismatch repair protein	630_gi|115251022	**99.69**	**0**	**99.84**	**0**	**99.53**	**0**
804	DeoR-family transcriptional regulator	630_gi|115251165	**99.61**	**1.00E-132**	**100**	**4.00E-133**	**100**	**3.00E-133**
805	putative phosphomannomutase/phosphoglycerate mutase	630_gi|115251833	**99.82**	**0**	**99.65**	**0**	**99.65**	**0**
806	putative DNA repair protein (nucleotide pyrophosphatase)	630_gi|115252458	**100**	**4.00E-79**	**100**	**4.00E-76**	**100**	**4.00E-79**
807	excinuclease ABC subunit B	630_gi|115252472	**100**	**0**	**100**	**0**	**100**	**0**

808	nusA, 1689456-1690610 (Clockwise) Hypothetical Protein nusA transcription elongation protein	sta_BASYS01606 (115250342)	**100**	**0**	**99.48**	**0**	**99.48**	**0**
809	BASYS00420, 401655-400888 (CounterClockwise) Hypothetical Protein BASYS00420 hypothetical	stb_BASYS00420	87.06	7.00E-121	100	2.00E-135		
812	deaD, 915268-916881 (Clockwise) Hypothetical Protein dead putative ATP-dependent RNA helicase	stt_BASYS00904 (115249778)	**100**	**0**	**100**	**0**	**100**	**0**
814	putative ferrous iron transport protein A	630_gi|115250789	**100**	**6.00E-37**	**100**	**7.00E-37**	**100**	**6.00E-37**
864	prolyl-tRNA synthetase	630_gi|115249052	**99.82**	**0**	**100**	**0**	**100**	**0**
865	transcriptional antiterminator	630_gi|115251564	**100**	**2.00E-157**	**100**	**2.00E-157**	**100**	**2.00E-157**
867	putative DNA polymerase III, delta' subunit	630_gi|115252610	**100**	**3.00E-176**	**100**	**3.00E-176**	**100**	**3.00E-176**
868	yjeE, 148838-149293 (Clockwise) Hypothetical Protein yjeE putative ATP/GTP hydrolase	stt_BASYS00152 (115249158)	**100**	**1.00E-83**	**100**	**1.00E-83**	**100**	**1.00E-83**
874	ABC transporter, ATP-binding protein	630_gi|115250659	**100**	**5.00E-138**	**100**	**5.00E-138**	**100**	**5.00E-138**
876	BASYS01324, 1412801-1414714 (Clockwise) Hypothetical Protein BASYS01324 hypothetical	sta_BASYS01324 (115250058)	**100**	**0**	**98.74**	**0**	**99.22**	**0**
878	phosphoglycerate kinase	630_gi|115252230	**99.75**	**0**	**100**	**0**	**100**	**0**
883	conserved hypothetical protein	630_gi|115252219	**99.39**	**2.00E-92**	**99.39**	**2.00E-92**	**98.77**	**5.00E-92**
886	zntA, 695882-698269 (Clockwise) Hypothetical Protein zntA putative heavy-metal-transporting ATPase	sta_BASYS00689 (115249322)	**100**	**0**	**98.99**	**0**	**98.49**	**0**
888	chromosomal replication initiator protein	630_gi|115249004	**100**	**0**	**100**	**0**	**100**	**0**
890	ompR, 2456141-2455449 (CounterClockwise) Hypothetical Protein ompR two-component system response regulator	stt_BASYS02359 (115251341)	**100**	**4.00E-129**	**100**	**4.00E-129**	**100**	**4.00E-129**
893	conserved hypothetical protein	630_gi|115251465	**100**	**0**	**100**	**0**	**100**	**0**
896	putative aminotransferase	630_gi|115251747	**99.24**	**0**	**99.75**	**0**	**100**	**0**
901	BASYS03068, 3231844-3230237 (CounterClockwise) Hypothetical Protein BASYS03068 No match by BLAST	stb_BASYS03068			100	0		
902	putative flagellar biosynthesis protein	630_gi|115249238	**92.03**	**1.00E-68**	**92.03**	**2.00E-68**	**92.03**	**2.00E-68**
904	putative helicase	630_gi|115250059	**99.59**	**0**	**99.86**	**0**	**100**	**0**

906	putative lipoprotein	630_gi|115250384	**100**	**2.00E-84**	**100**	**2.00E-84**	**100**	**2.00E-84**
908	630_gi|115249949|emb|CAJ67769.1|-R	630_gi|115249949						
909	putative phage-related protein	630_gi|115250413	97.44	2.00E-16	100	1.00E-16		
910	conserved hypothetical protein	630_gi|115249410	**99.07**	**1.00E-58**	**100**	**3.00E-60**	**100**	**3.00E-60**
911	BASYS00592, 612382-612774 (Clockwise) Hypothetical Protein BASYS00592 Hypothetical protein (Flagellar assembly factor FliW)	sta_BASYS00592 115249241	**100**	**8.00E-71**	**99.23**	**5.00E-70**	**99.23**	**5.00E-70**
915	recT, 153449-152616 (CounterClockwise) Hypothetical Protein recT recombination and repair protein RecT [*Clostridium difficile* QCD-63q42]	stb_BASYS00175			100	5.00E-164		
921	putative esterase	630_gi|115250430	**99.3**	**0**	**99.65**	**0**	**99.12**	**0**

Table 3.6 *– Proteins identified in strain Tra5/5 only showing BLAST percentage identification and e-value scores to each of the three strains.*

	Strain Tra5/5 only		Strain A		Strain B-1		Strain Tra5/5	
			% ID	E value	% ID	E value	% ID	E value
295	seryl-tRNA synthetase	630_gi|115249017	**99.53**	**0**	**100**	**0**	**100**	**0**
316	fdhF, 2319506-2317329 (CounterClockwise) Hypothetical Protein fdhF (putative formate dehydrogenase)	stt_BASYS02242 (115251232)	**98.74**	**0**	**99.58**	**0**	**100**	**0**
360	yiaJ, 3855532-3856284 (Clockwise) Hypothetical Protein yiaJ	stt_BASYS03597	31.84	3.00E-25	31.84	3.00E-25	100	4.00E-143
362	putative molybdenum-binding subunit of oxidoreductase	630_gi|115251153	**99.73**	**0**	**99.87**	**0**	**100**	**0**
378	putative electron transfer protein	630_gi|115252303	**100**	**0**	**100**	**0**	**100**	**0**
380	asparaginyl-tRNA synthetase	630_gi|115251299	**100**	**0**	**100**	**0**	**100**	**0**
451	cysS, 393517-394923 (Clockwise) Hypothetical Protein cysS cysteinyl-tRNA synthetase	sta_BASYS00384 (115249055)	**100**	**0**	**99.57**	**0**	**99.79**	**0**
481	putative chorismate biosynthesis-related protein includes: phospho-2-dehydro-3-deoxyheptonate aldolase and chorismate mutase	630_gi|115250878	**99.7**	**0**	**100**	**0**	**100**	**0**
527	putative FAD-binding subunit of oxidoreductase	630_gi|115251155	**100**	**4.00E-151**	**99.62**	**8.00E-151**	**100**	**4.00E-151**
539	dihydroorotate dehydrogenase, catalytic subunit	630_gi|115249196	**99.33**	**7.00E-164**	**99**	**2.00E-163**	**99.33**	**7.00E-164**
542	putative aminopeptidase	630_gi|115250096	**98.92**	**0**	**99.35**	**0**	**99.78**	**0**
547	threonine synthase	630_gi|115251172	**99.39**	**0**	**99.8**	**0**	**99.59**	**0**
572	putative radical SAM family protein	630_gi|115251801	**99.56**	**0**	**99.56**	**0**	**99.56**	**0**
591	phosphoglucomutase/phospho mannomutase mutase	630_gi|115249128	**99.79**	**0**	**99.57**	**0**	**99.57**	**0**
592	conserved hypothetical protein	630_gi|115250761	**99.72**	**0**	**100**	**0**	**100**	**0**
593	N-carbamoyl-L-amino acid hydrolase	630_gi|115251081	**99.75**	**0**	**99.51**	**0**	**99.75**	**0**
594	putative thymidylate kinase	630_gi|115252611	**100**	**5.00E-134**	**100**	**5.00E-134**	**100**	**5.00E-134**
595	pyrroline-5-carboxylate reductase	630_gi|115252337	**99.63**	**4.00E-153**	**99.63**	**8.00E-153**	**100**	**2.00E-153**
606	chorismate synthase (5-enolpyruvylshikimate-3-phosphate phospholyase)	630_gi|115250881	**99.72**	**0**	**99.44**	**0**	**99.72**	**0**
658	putative ATP/GTP-binding protein	630_gi|115249805	**99.06**	**0**	**99.53**	**0**	**100**	**0**
690	P-protein includes: chorismate mutase and prephenate dehydratase	630_gi|115250882	**97.99**	**0**	**98.74**	**0**	**99.5**	**0**
691	UDP-N-acetylmuramate--L-alanine ligase	630_gi|115252579	**100**	**0**	**100**	**0**	**100**	**0**

707	dihydrodipicolinate reductase	630_gi|115252283	**99.6**	**3.00E-135**	**100**	**1.00E-135**	**100**	**1.00E-135**
728	radical SAM-superfamily protein	630_gi|115250362	**99.55**	**0**	**99.77**	**0**	**99.77**	**0**
729	glutamine synthetase	630_gi|115250379	**100**	**0**	**100**	**0**	**100**	**0**
730	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alan ine ligase	630_gi|115251706	**99.12**	**0**	**99.78**	**0**	**99.56**	**0**
731	putative GTP pyrophosphokinase	630_gi|115251796	**100**	**0**	**100**	**0**	**100**	**0**
732	galU, 3885914-3884934 (CounterClockwise) Hypothetical Protein galU	stt_BASYS03633	**99.69**	**6.00E-179**	**99.69**	**7.00E-179**	**100**	**3.00E-179**
760	oppF, 1277546-1278490 (Clockwise) Hypothetical Protein oppF ABC-type oligopeptide transport system, ARPase component	sta_BASYS01208	**100**	**4.00E-179**	**100**	**4.00E-179**	**99.36**	**1.00E-177**
791	putative 2-nitropropane dioxygenase	630_gi|115251174	**100**	**2.00E-180**	**100**	**2.00E-180**	**100**	**2.00E-180**
792	potG, 466136-466891 (Clockwise) Hypothetical Protein potG ABC transporter, ATP-binding protein	stb_BASYS00509 (115249891)	**99.24**	**2.00E-60**	**100**	**3.00E-119**	**99.49**	**5.00E-87**
795	putative DNA mismatch repair protein	630_gi|115249725	**99.49**	**0**	**99.75**	**0**	**99.75**	**0**
822	DNA repair protein	630_gi|115249030	**99.78**	**0**	**100**	**0**	**100**	**0**
823	putative hydantoinase	630_gi|115250762	**100**	**0**	**99.73**	**0**	**99.74**	**0**
824	PhoH-like protein	630_gi|115251494	**99.41**	**4.00E-174**	**99.7**	**4.00E-174**	**99.7**	**4.00E-174**
825	cinserved hypothetical protein	630_gi|115251677	**99.56**	**0**	**99.78**	**0**	**99.56**	**0**
826	probable peptidase	630_gi|115251712	**100**	**0**	**100**	**0**	**100**	**0**
829	yibO, 3925985-3924453 (CounterClockwise) Hypothetical Protein yibO 2,3-bisphosphoglycerate-independent phosphoglycerate mutase	sta_BASYS03659 (115252228)	**100**	**0**	**99.41**	**0**	**99.61**	**0**
830	purF, 231027-232394 (Clockwise) Hypothetical Protein purF amidophosphoribosyltransferase	stt_BASYS00227 (115249228)	**100**	**0**	**100**	**0**	**100**	**0**
866	hydrogenase	630_gi|115252370	**100**	**0**	**100**	**0**	**100**	**0**
873	conserved hypothetical protein	630_gi|115250624	**100**	**1.00E-47**	**100**	**1.00E-47**	**100**	**1.00E-47**
877	hsdM, 3719141-3716964 (CounterClockwise) Hypothetical Protein hsdM type I restriction-modification system specificity subunit	stt_BASYS03486	24.63	1.00E-07	24.93	3.00E-08	100	0
882	cyclopropane-fatty-acyl-phospholipid synthase	630_gi|115249187	**99.75**	**0**	**99.49**	**0**	**99.49**	**0**

895	putative deoxyribonuclease	630_gi|115252600	**99.61**	**9.00E-151**	**100**	**2.00E-151**	**99.61**	**3.00E-151**
925	cell surface protein	630_gi|115251246	**100**	**0**	**100**	**0**	**100**	**0**

3.2.3 – Protein Identification III – DIGE Analysis

In order to accurately compare the 2D profiles of the different strains, DIGE analysis was used. Two biological replicates of the standard protein extracts from the three strains were labelled with fluorescent CyDyes, and run alongside a pooled protein standard containing a mixture of all protein samples in the experiment, as shown in *Table 3.7* below.

***Table 3.7** – Experimental setup and sample labelling for the 2D DIGE experiment.*

Gel	Dye		
	Cy2	**Cy3**	**Cy5**
1	Pooled Standard	Strain B-1 rep1	Strain A rep1
2	Pooled Standard	Strain A rep1	Strain Tra5/5 rep1
3	Pooled Standard	Strain Tra5/5 rep1	Strain B-1 rep1
4	Pooled Standard	Strain B-1 rep2	Strain A rep2
5	Pooled Standard	Strain A rep2	Strain Tra5/5 rep2
6	Pooled Standard	Strain Tra5/5 rep2	Strain B-1 rep2

Gels were imaged immediately after electrophoresis, and the images from each gel in the experiment shown in *Figure 3.13* below.

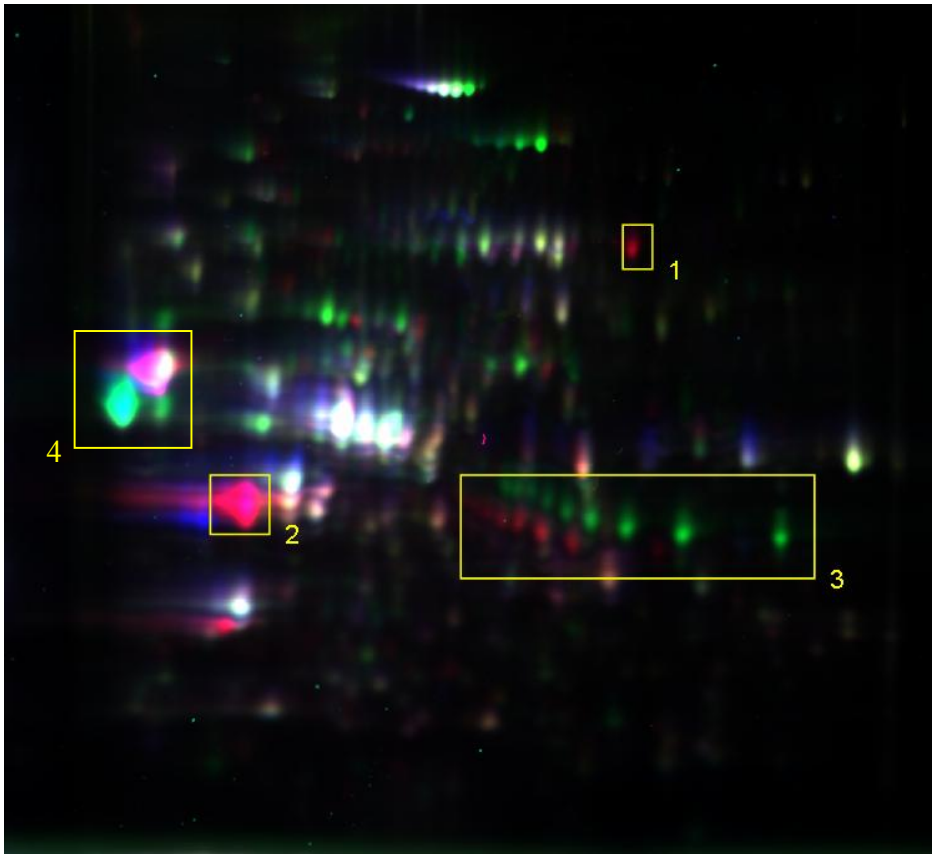


Figure 3.13(i) *– DIGE image gel 1. The Cy2 channel (blue) is the internal standard containing a mix of protein extracts from all strains. The Cy3 channel (green) is a protein extract from strain B-1(biological rep 1), and the Cy5 channel (red) a protein extract from strain A (biological rep 1). Spots appear white where all channels overlap. Examples of differentially expressed proteins are shown in the yellow boxes, with a surface protein gi\|115251884 in box 1, the LMW SLP in box 2, flagellin in box 3 and the HMW SLP shown in box 4. The surface protein and the LMW SLP are clearly present in the A strain (red) but not the B-1 strain (green). The HMW SLP and the series of flagellin spots are clearly migrating differently in the two strains.*

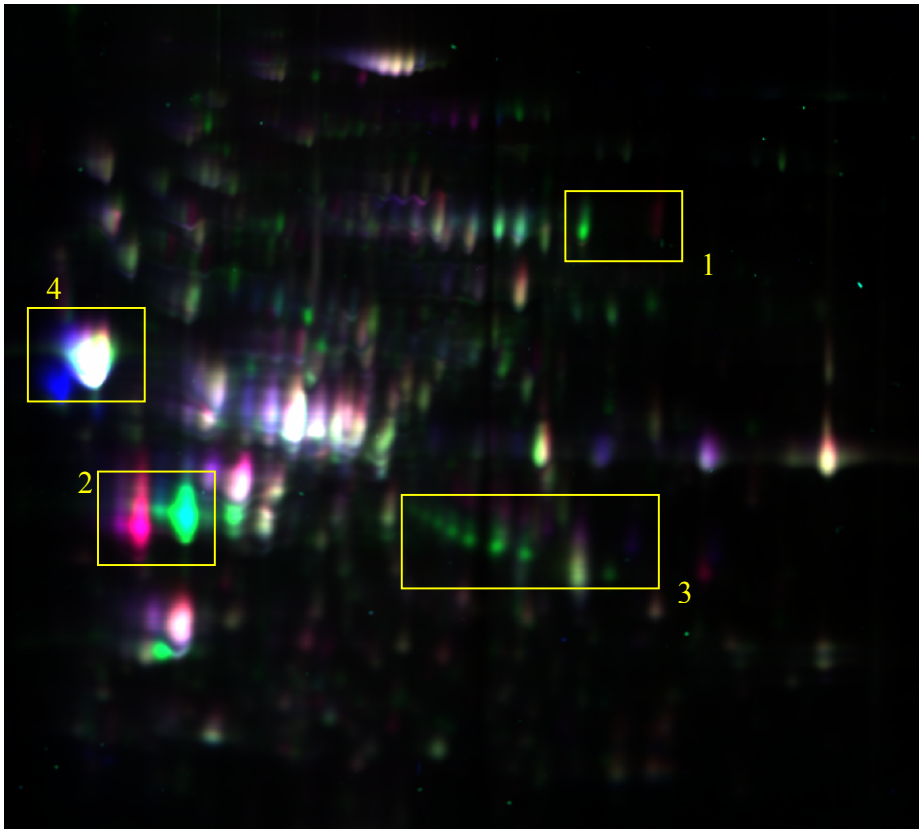


Figure 3.13(ii) *– DIGE image gel 2. The Cy2 channel (blue) is the internal standard containing a mix of protein extracts from all strains. The Cy3 channel (green) is a protein extract from strain A (biological rep 1), and the Cy5 channel (red) a protein extract from strain Tra5/5 (biological rep 1). Spots appear white where all channels overlap. Examples of differentially expressed proteins are shown in the yellow boxes as described in Figure 3.13(i). The surface protein in box 1 and the LMW-SLP (box 2) show clear migration differences between the A and Tra5/5 strains, whereas the HMW SLP (box 4) does not.*

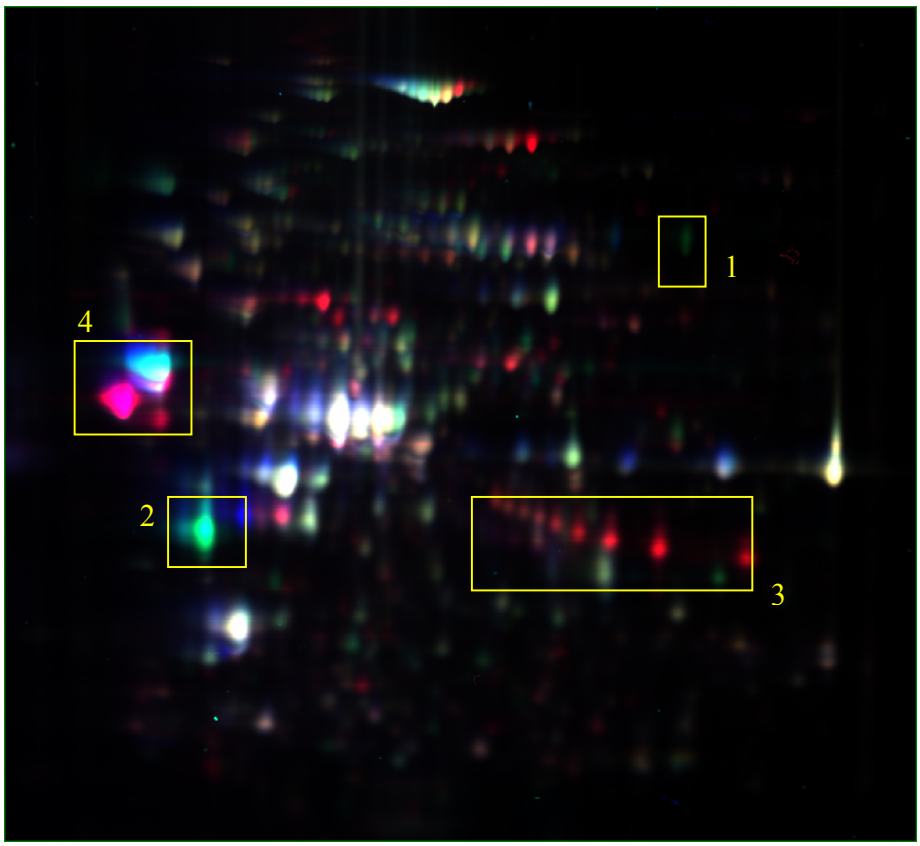


Figure 3.13(iii) *– DIGE image gel 3. The Cy2 channel (blue) is the internal standard containing a mix of protein extracts from all strains. The Cy3 channel (green) is a protein extract from strain Tra5/5 (biological rep 1), and the Cy5 channel (red) a protein extract from strain B-1 (biological rep 1). Spots appear white where all channels overlap. Examples of differentially expressed proteins are shown in the yellow boxes as described in Figure 3.13(i). The surface protein and LMW SLP (box 1 and 2) are present in strain Tra5/5 (green) but not strain B-1 (red) and the HMW SLP (box 4) clearly shows migration differences between the strains. The series of flagellin spots (box 3) can be clearly seen for strain B-1 but not Tra5/5.*

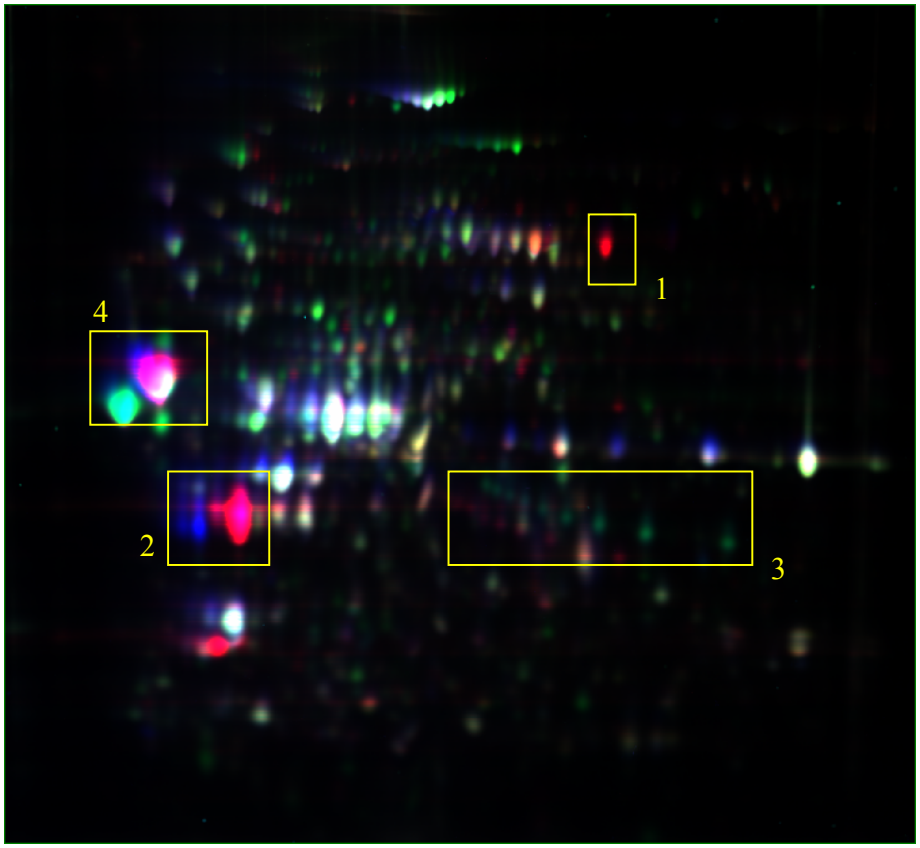


Figure 3.13(iv) *– DIGE image gel 4. The Cy2 channel (blue) is the internal standard containing a mix of protein extracts from all strains. The Cy3 channel (green) is a protein extract from strain B-1 (biological rep 2), and the Cy5 channel (red) a protein extract from strain A (biological rep 2). Spots appear white where all channels overlap. Examples of differentially expressed proteins are shown in the yellow boxes as described in Figure 3.13(i). The same variations in protein expression can be seen as for gel 1 (Figure 3.13(i)).*

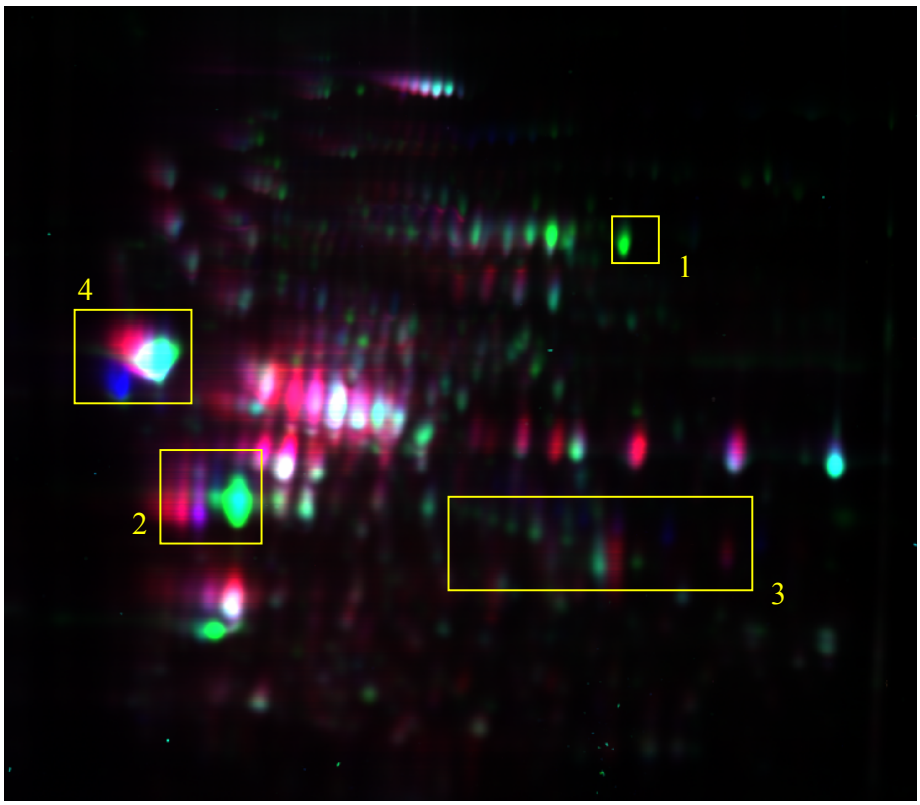


Figure 3.13(v) *– DIGE image gel 5. The Cy2 channel (blue) is the internal standard containing a mix of protein extracts from all strains. The Cy3 channel (green) is a protein extract from strain A (biological rep 2), and the Cy5 channel (red) a protein extract from strain Tra5/5 (biological rep 2). Examples of differentially expressed proteins are shown in the yellow boxes as described in Figure 3.13(i). The same expression differences can be seen as for gel 2 (Figure 3.13(ii)).*

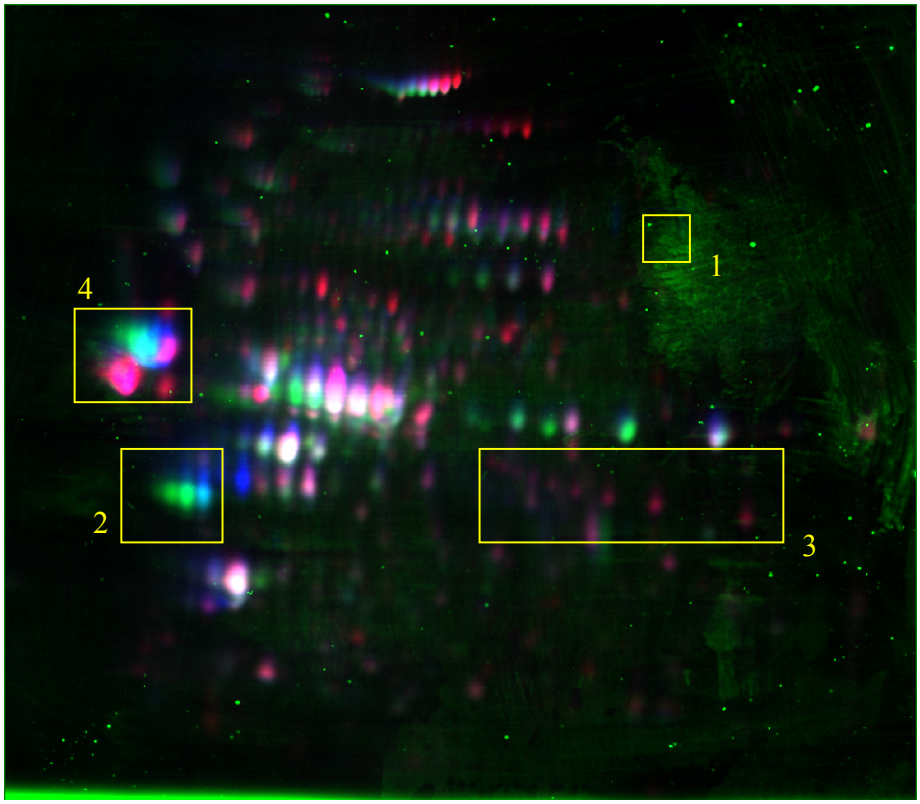


Figure 3.13(vi) *– DIGE image gel 6. The Cy2 channel (blue) is the internal standard containing a mix of protein extracts from all strains. The Cy3 channel (green) is a protein extract from strain Tra5/5 (biological rep 2), and the Cy5 channel (red) a protein extract from strain B-1 (biological rep 2). Spots appear white where all channels overlap. Examples of differentially expressed proteins are shown in the yellow boxes as described in Figure 3.13(i). The surface protein (box 1) is not visible in either strain, but other migration differences are the same as those for gel 3 (Figure 3.13(iii)).*

Yellow box 1 highlights an S-layer protein (gi|126700407) with a 43.7 fold greater concentration in strain A, and a normalised volume of 5.4 compared to 0.1 for strain B-1 and 0.5 for strain Tra5/5. The same protein was also identified as a different spot, up-regulated in strain Tra5/5 with a 33.4 fold greater concentration in strain A, and normalised volumes of 2.5 for strain Tra5/5

compared to 0.1 for strain B-1 and 0.8 for strain A. This protein was identified on the reference maps of strain A and Tra5/5 (spot J12, *Figures 3.8 and 3.10*) but not B-1.

The yellow boxes, 2 and 4 show the S-layer protein most similar to the SlpA of the 630 reference strain (gi|126700409), corresponding to the mature high molecular weight SLP and the low molecular weight SLP. In both boxes, this protein has migrated differently between strains, due to the difference in molecular weight and pI arising from the large variation in amino acid sequence. On the reference maps, this protein was identified as one spot in the profile of strain B-1 (spot A1, *Figure 3.9*) but two spots in strains A and Tra5/5 (spots A1 and F6 *Figures 3.8 and 3.10*). Yellow box 2 highlights this second spot, F6. The spot shows a 19.1 fold greater concentration in strain A, and a normalised volume of 3.5 compared to 0.2 for strain B-1 and 0.2 for strain Tra5/5. The equivalent spot for the Tra5/5 version of this LMW SLP shows a 9.6 fold greater concentration in strain Tra5/5 and a normalised volume of 1.3 compared to 0.1 in strain B-1 and 0.3 in strain A.

Yellow box 3 highlights the differences in the modified flagellin proteins of the B-1 and A strains, where the series of Cy3 and Cy5 labelled proteins are clearly visible as separate spots which have migrated differently within the gel. This suggests that the modification process is different between the strains.

The images from all six gels were aligned, and statistical concentration differences between different strains analysed using the SameSpots software (Progenesis). A total of 453 proteins were matched across the standards of all

six gels. Proteins with an anova value of $p<0.05$ and a greater than ±2-fold difference were considered to differ significantly between the strains. Correlation analysis was used on the 112 spots which met these criteria in order to compare expression profiles between strains. As shown below, 28 proteins showed a significantly greater concentration in the B-1 strain, a different 28 showed significantly greater concentration in the A strain and 21 proteins showed a significantly greater concentration in the Tra5/5 strain. The DIGE images were then matched to the picking gels used for the reference maps (*Figures 3.8, 3.9 and 3.*10) to identify these proteins. The ranks, anova scores, fold increase and normalised volumes of these significant proteins are shown in *Appendix A4*.

Of the 28 spots with significantly higher concentration in the A strain ($p<0.05$, fold difference >±2), 15 (54%) matched to spots on the reference map (*Figure 3.14*). Ten of these had been identified by MASCOT, as five different proteins (*Table 3.8*), leading to identification of 36% of the significant spots. Of these spots of interest, the S-layer protein spots J12 (*Figure 3.8*, *Table 3.8*, sta_BASYS03267) and F6 (*Figure 3.8*, *Table 3.8*, sta_BASYS03269) showed the highest fold differences of 43.7 and 19.1 respectively. Five different spots (marked as M2, M4-M7 in *Figure 3.8*) were identified as flagellin, with fold differences ranging from 6.4 to 4.3, and an ABC transporter protein (marked as F5 in *Figure 3.8*) was identified with a lower fold difference of 2.5.

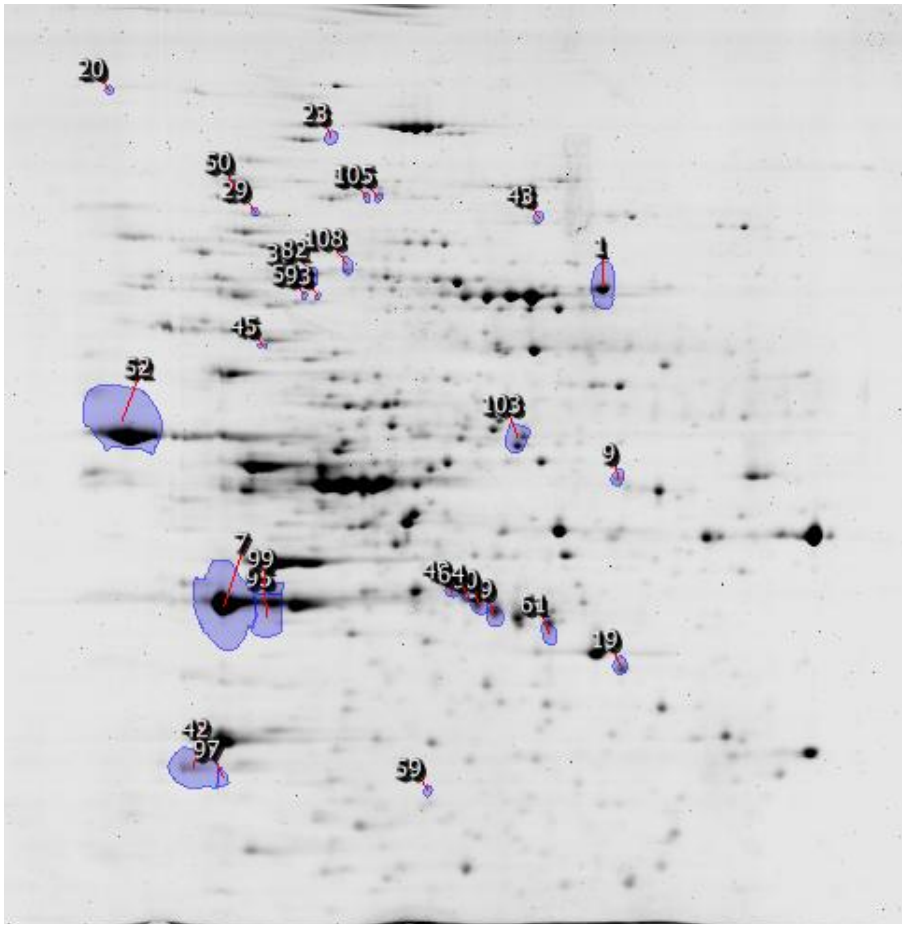


Figure 3.14 *– Protein spots with statistically significantly higher concentration in strain A. The proteins identified as up-regulated in strain A by correlation analysis were matched to the 'picking gel' used to create the strain A reference map. The numbers indicate the rank of the protein, with protein 1 showing the greatest fold difference between the strains (Appendix A4)*

Table 3.8 *– Protein spots with significantly higher concentration in strain A*

Reference map ID	Protein	ID	Rank
A1 and F6	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 (Cell Surface Protein)	sta_BASYS03269_	52 and 7 respectively
F5	ABC transporter, substrate-binding lipoprotein	630_gi|115249887|emb|CAJ67706.1|_	95
J12	BASYS03267, 3473846-3471975 (CounterClockwise) Hypothetical Protein BASYS03267 (cell surface protein (putative S-layer protein precursor))	sta_BASYS03267_	1
L5	glyA, 3392427-3391183 (CounterClockwise) Hypothetical Protein glyA (putative serine hydroxymethyltransferase)	sta_BASYS03201_	103
M2 and M4-M7	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	sta_BASYS00598_	61, 39, 60, 64 and 49 respectively

Of the 28 protein spots with significantly higher concentration in the B-1 strain ($p<0.05$, fold difference $\geq\pm2$), 17 (71%) were matched to spots on the B-1 reference map (*Figure 3.15*). All but three of these had been identified by MASCOT, as ten different proteins (*Table 3.9*), leading to identification of 61% of the significant spots. Of these spots, the greatest fold difference of 22.6 was seen in the S-layer protein spot A1 (*Figure 3.9, Table 3.9*, stb_BASYS03072). Four flagellin spots had a significantly higher concentration in strain B-1, with the spots marked M1 and M3 on *Figure 3.9* showing a 21.1 and 22.4 fold increase respectively. Three spots (marked as D5-D7 on *Figure 3.9*), identified as ABC transporter binding proteins, showed fold differences of 4.6, 6.7 and 6.6. In addition, a number of metabolic proteins were up-regulated in strain B-1, including a pyruvate-flavodoxin oxidoreductase (spots B1 and B2, *Figure 3.9*)

with fold differences of 4.2 and 6.7 and an alcohol dehydrogenase (spot I2, *Figure 3.9*) with a large fold difference of 12.2.

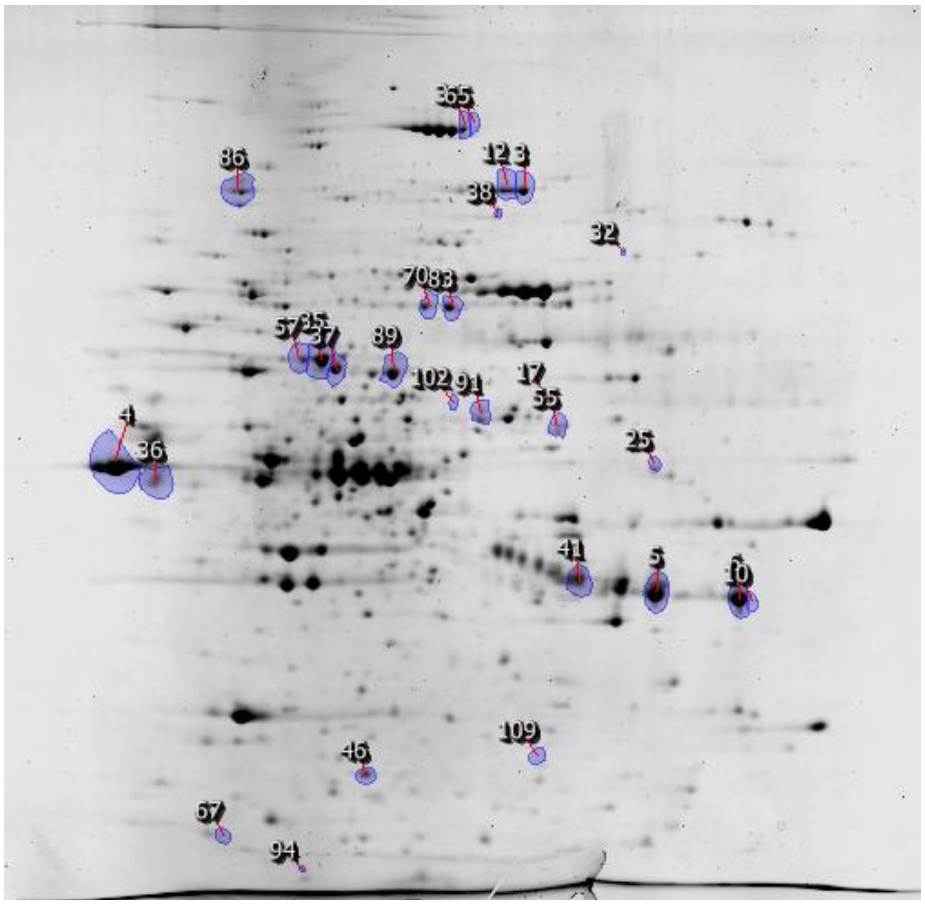


Figure 3.15 *– Protein spots with statistically significantly higher concentration in strain B-1. The proteins identified as up-regulated in strain B-1 by correlation analysis were matched to the 'picking gel' used to create the strain B-1 reference map. The numbers indicate the rank of the protein, with protein 1 showing the greatest fold difference between the strains (Appendix A4).*

Table 3.9 *– Protein spots with significantly higher concentration in strain B-1. Legend – (x2) indicates that one spot picked from the reference map was identified as two spots by the SameSpots image analysis software.*

Reference map ID	Protein	ID	Rank
A1 and A8	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	4 and 36 respectively
B1 and B2	pyruvate-flavodoxin oxidoreductase	630_gi|115251733|emb|CAJ69568.1|	65 and 34 respectively
D5 and D6	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872|emb|CAJ67689.1|	57and 35 respectively
D7	oligopeptide ABC transporter, substrate-binding protein	630_gi|115251723|emb|CAJ69558.1|	37
D9	glycine/sarcosine/betaine reductase complex component C beta subunit	630_gi|115251404|emb|CAJ69236.1|	89
G10	glycine reductase complex component B alpha and beta subunits	630_gi|115251407|emb|CAJ69239.1|	46
I2	adhE, 3383815-3381161 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase)	stb_BASYS03189	12
J7 and J8	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184|emb|CAJ66996.1|	83 and 70 respectively
K5	radical SAM-superfamily protein	630_gi|115251209|emb|CAJ69040.1|	91
M1(x2), M3 and M5	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	6, 10, 5 and 41 respectively

Twenty-one spots showed a significantly higher concentration in strain Tra5/5 ($p<0.05$, fold difference $>\pm 2$), only 48% of which matched to spots on the Tra5/5 reference map (*Figure 3.16*). All of these had been identified using MASCOT (*Table 3.10)*, as eight different proteins. Again, the largest fold difference of 33.4 was seen in an S-layer protein, spot J12 (*Figure 3.10, Table 3.10,* stt_BASYS02896).

Three spots from the DIGE images were matched to an additional S-layer protein marked spot F6 on *Figure 3.10* (*Table 3.10,* stt_BASYS00983) with high fold differences of 10.4, 9.6 and 8.8. Similarly to strain B-1, some metabolic proteins showed a higher concentration in strain Tra5/5 including an oxidoreductase (marked as spot B5 on *Figure 3.10*) with a 2.4 fold difference, and a dehydrogenase (marked as spot L10 on *Figure 3.10*) identified as two spots with fold differences of 3.5 and 3.8.

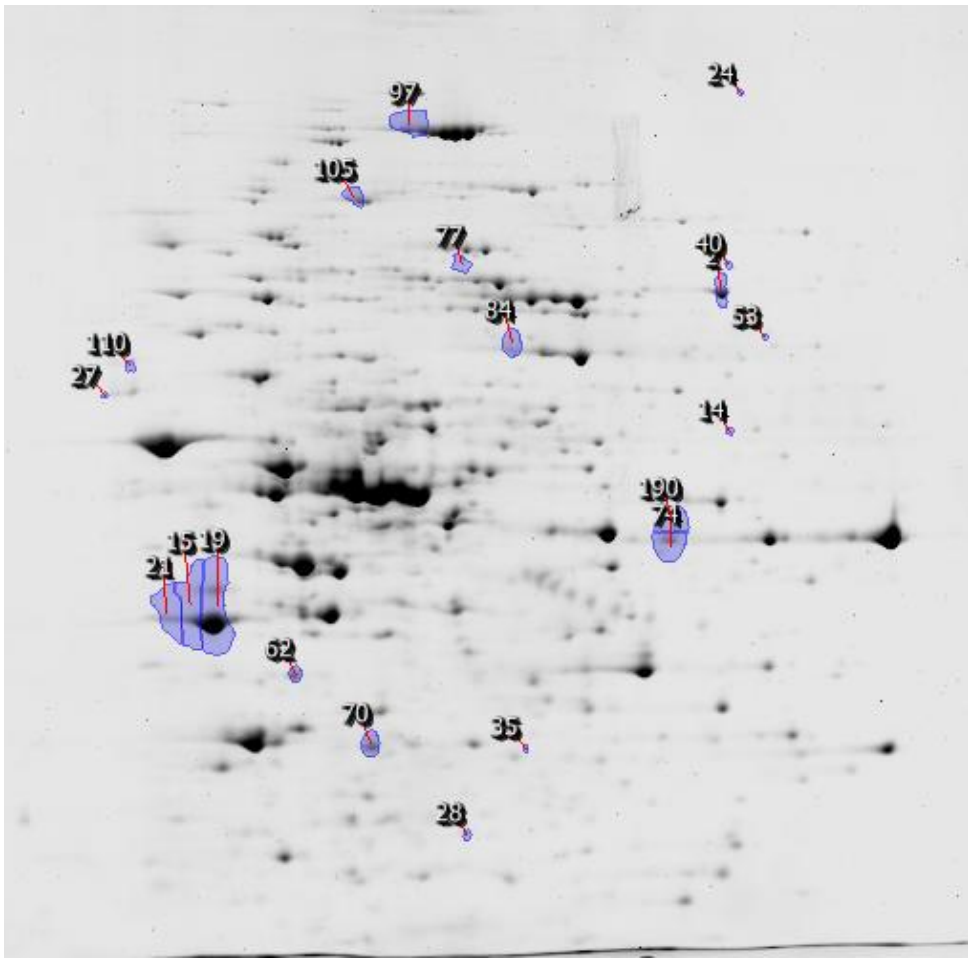


Figure 3.16 *– Protein spots with statistically significantly higher concentration in strain Tra5/5.*

The proteins identified as up-regulated in strain Tra5/5 by correlation analysis were matched to the 'picking gel' used to create the strain Tra5/5 reference map. The numbers indicate the rank of the protein, with protein 1 showing the greatest fold difference between the strains (Appendix A4).

Table 3.10 *– Protein spots with significantly higher concentration in the Tra5/5 strain. Legend – (x2) or (x3) indicates that one spot picked from the reference map was identified as two or three spots by the SameSpots image analysis software. Two protein identifications indicate that one spot on the reference map was identified as two different proteins by MS.*

Reference map ID	Protein	ID	Rank
B5	ydbK, 2906795-2903256 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase2	stt_BASYS02779	97
F6 (x3)	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898 (Cell Surface protein)	stt_BASYS02898	19, 15 and 21 respectively
F12	ycgM, 996287-997189 (Clockwise) Hypothetical Protein ycgM (putative hydrolase)	stt_BASYS00983	62
G8	BASYS01847, 1900865-1901566 (Clockwise) Cytoplasmic Protein (conserved hypothetical protein)	stt_BASYS01847	70
I5	leuS, 2716913-2714493 (CounterClockwise) Hypothetical Protein leuS (leuS, 3140547-3138127 (CounterClockwise) Hypothetical Protein leuS) (leucyl-tRNA synthetase)	stt_BASYS02608	105
J12	BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (cell surface protein (putative S-layer protein precursor))	stt_BASYS02896	2
L10 (x2)	caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	stt_BASYS00464	74 and 190 respectively
	cfa, 188449-187262 (CounterClockwise) Hypothetical Protein cfa (cyclopropane-fatty-acyl-phospholipid synthase)	stt_BASYS00186	

During this analysis, different spots showing significant differences between strains by the correlation analysis were often identified as the same protein on the reference maps, such as the flagellin protein (BASYS02546) in strain B-1 which matched to four different spots with significantly higher concentrations in strain B-1, and the S-layer protein F6 (gi|126700409) in strain Tra5/5 which matched to three different spots that showed significantly higher concentration in strain Tra5/5.

A total of five different up-regulated proteins were identified in strain A, 12 in strain B-1 and eight in strain Tra5/5. In order to confirm that the difference in concentration of these up-regulated proteins was due to regulatory rather than genomic differences, the presence of the corresponding genes in all three genomes was verified. The amino acid sequences of the identified proteins were therefore matched to *in silico* gene products from the unfinished genomes using BLAST searching. Each protein sequence was matched against the gene products of each of the three genomes in turn to confirm the presence of the corresponding gene in the genome. Percentage identification and e-values for each match are shown in *Table 3.11* below.

Table 3.11 *– Percentage identification scores and e-values for BLAST comparison of the sequences of differentially expressed proteins with each of the three genomes. Scores highlighted in bold are considered significant and indicate the presence of the corresponding gene in the genome. Lower scores indicate considerable sequence variation between strains.*

Upregulated proteins		**Strain A genome**		**Strain B-1 genome**		**Strain Tra5/5 genome**	
Strain	Protein	%id	e	%id	e	%id	e
	A1, F6 - sta_BASYS03269 BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 (Cell Surface Protein)	**100**	**0**	52.71	4.00E -106	**88.71**	**0**
	F5 - 630_gi|15249887 ABC transporter, substrate-binding lipoprotein	**-**	**-**	**99.71**	**0**	**99.71**	**0**
	J12 - sta_BASYS03267 BASYS03267, 3473846-3471975 (CounterClockwise) Hypothetical Protein BASYS03267 (cell surface protein (putative S-layer protein precursor))	**100**	**0**	36.3	3.00E -79	**98.07**	**0**
	L5 - sta_BASYS03201 glyA, 3392427-3391183 (CounterClockwise) Hypothetical Protein glyA (putative serine hydroxymethyltransferase)	**100**	**0**	**99.52**	**0**	**99.52**	**0**
A	**M2-M7 -** sta_BASYS00598 fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	**100**	**6.00E -169**	**99.37**	**2.00E -168**	**99.37**	**2.00E -168**
	A1, A8 - stb_BASYS03072 (S-layer) BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	52.71	6.00E -113	**100**	**0**	52.91	2.00E -113
	B1, B2 - 630_gi|11521733 pyruvate-flavodoxin oxidoreductase	**100**	**0**	**99.41**	**0**	**99.58**	**0**
B-1	**D5, D6 -** 630_gi|115249872 oligopeptide ABC transporter, substrate-binding lipoprotein	**99.8**	**0**	**100**	**0**	**100**	**0**

	D7 - 630_gi|115251723 oligopeptide ABC transporter, substrate-binding protein	**99.61**	**0**	**99.42**	**0**	**99.81**	**100**
	D9 - 630_gi|115251404 glycine/sarcosine/betaine reductase complex component C beta subunit	**99.8**	**0**	**100**	**0**	**100**	**0**
	G10 - 630_gi|115251407 glycine reductase complex component B alpha and beta subunits	**100**	**0**	**100**	**0**	**100**	**0**
	I2 - stb_BASYS03189 adhE, 3383815-3381161 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase)	**99.21**	**0**	**100**	**0**	**99.89**	**0**
	J7, J8 - 630_gi|115249184 putative oxidoreductase, acetyl-CoA synthase subunit	**100**	**0**	**100**	**0**	**100**	**0**
	K5 - 630_gi|115251209 radical SAM-superfamily protein	**100**	**0**	**100**	**0**	**100**	**0**
B-1	**M1-M5 –** stb_BASYS00368 fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	**99.37**	**9.00E -180**	**100**	**1.00E -180**	**100**	**1.00E -180**
	B5 – stt_BASYS02779 ydbK, 2906795-2903256 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase2	**99.58**	**0**	**99.83**	**0**	**100**	**0**
	F6 – stt_BASYS02898 BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898 (Cell Surface protein)	**88.71**	**0**	52.91	1.00E -106	**100**	**0**
	F12 – stt_BASYS00983 ycgM, 996287-997189 (Clockwise) Hypothetical Protein ycgM (putative hydrolase)	**99.67**	**2.00E -174**	**99.33**	**6.00E -174**	**100**	**1.00E -175**
Tra 5/5	**G8 –** stt_BASYS01847 BASYS01847, 1900865-1901566 (Clockwise) Cytoplasmic Protein (conserved hypothetical protein)	**100**	**2.00E -134**	**100**	**2.00E -134**	**100**	**2.00E -134**

	I5 – stt_BASYS02608 leuS, 2716913-2714493 (CounterClockwise) Hypothetical Protein leuS (leuS, 3140547-3138127 (CounterClockwise) Hypothetical Protein leuS) (leucyl-tRNA synthetase)	**99.88**	**0**	**100**	**0**	**100**	**0**
	J12 – stt_BASYS02896 BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (cell surface protein (putative S-layer protein precursor))	**98.07**	**0**	36.6	3.00E -80	**100**	**0**
	L10 – stt_BASYS00464 caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	**100**	**0**	**100**	**0**	**100**	**0**
Tra5/5	**L10 –** stt_BASYS00186 cfa, 188449-187262 (CounterClockwise) Hypothetical Protein cfa (cyclopropane-fatty-acyl-phospholipid synthase)	**99.24**	**0**	**100**	**0**	**100**	**0**

Of all the proteins deemed significant by the SameSpots software, the S-layer proteins were the only sequences which did not match with high homology to a predicted gene product from all of the three genomes.

Three of the six proteins with significantly higher concentration in strain A matched to predicted gene products in all three genomes with high homology, i.e. a percentage identification greater that 80% (scores highlighted in bold *Table 3.11*). However, the S-layer proteins BASYS01133 and BASYS01135, did not. Both these proteins showed high homology (89% and 99% respectively) to predicted gene products from strain Tra5/5, but much lower homology (53% and 36% respectively) to predicted gene products from strain B-1. There is an obvious anomaly in that the ABC transporter (F5) found to be unregulated in strain A matched with high homology to strains B-1 and Tra5/5 but not to A,

despite the fact the protein was detected in strain A. This ABC transporter seems to be conserved in a number of *C. difficile* strains, including similar 027 ribotype strains, so it seems unlikely that it would not be present in A. The individual peptides on which the identification was based were therefore searched against each of the genomes using stand-alone BLAST with the recommended parameters for short sequences. Again, the peptides all matched to the same protein sequence with high homology in strains B-1 and Tra5/5. For strain A, however, the peptides matched to a number of different proteins with low homology. The corresponding DNA sequence of this protein was then searched against the completed genome of strain A. Only 44 out of ~1000 nucleotides for the corresponding gene sequence were matched to the genome, indicating that the gene was not present in the genome. However, in the raw, un-annotated sequence data, the whole gene could be matched with high homology by BLAST. This indicates that when the sequences were reassembled using a newer version of the Roche assembly program 'Newbler' it didn't assemble some sequence data, which included this protein. This may be because it was a region considered to be of low sequence quality by the algorithm. The corresponding gene was nevertheless confirmed in all three strains.

Of the ten proteins with significantly higher concentration in strain B-1, nine matched to gene products from all three genomes with a high homology, but again, the S-layer protein did not. BASYS00187 showed a percentage identification of only 53% to gene products from both strains A and Tra5/5.

For strain Tra5/5, six of the eight significantly up-regulated proteins matched to gene products of all three genomes with high homology. The S-layer protein BASYS02999 showed high homology to strain A, with a percentage identification of 89%, a much lower percentage identification of 53% to strain B-1. BASYS02997 also showed high homology to strain A (98% identification) but a much lower homology to strain B-1 with a percentage identification score of only 37%.

3.3 – ProteoMiner Treatment

In order to increase the detection of low abundance proteins, and hence improve proteome coverage, protein extracts were treated with the ProteoMiner hexapeptide bead libraries. Two technical replicates of whole cell broth extracts were treated sequentially with an amino-terminus 'E-bead' hexapeptide library, and a carboxy-terminus 'SE-bead' library. Proteins were then eluted from the columns to give three different extracts; crude protein extracts, an E-bead eluate and an SE-bead eluate. These protein extracts were then compared to determine whether additional proteins were detected in the ProteoMiner treated samples.

3.3.1 Differences in 1D profiles

Firstly, the 1D gel profiles for treated and untreated samples were compared in order to assess the effects of the treatment (*Figure 3.17*).

The differences between the crude and ProteoMiner treated protein extracts were similar for each of the three strains. The extracts treated with the amino terminal E-bead library showed a similar profile to the crude samples, although some of the stronger bands, most notably that of ~55kDa, likely to correspond to

the highly abundant glutamate dehydrogenase, appear reduced in the E-bead treated samples. The two S-layer proteins of ~45 kDa and ~36 kDa are still, however, by far the most abundant proteins in the E-bead treated fractions, and their relative concentration does not appear to have been greatly altered by the treatment.

In contrast, the extracts treated by the carboxy-terminal SE-bead library have a markedly different 1D gel profile from the crude samples. The SE-bead treated extracts show a far smaller range in protein abundance, with the concentration of the S-layer proteins greatly reduced in comparison with both the crude and E-bead treated extracts. They also appear to show an increase in concentration of lower molecular weight proteins.

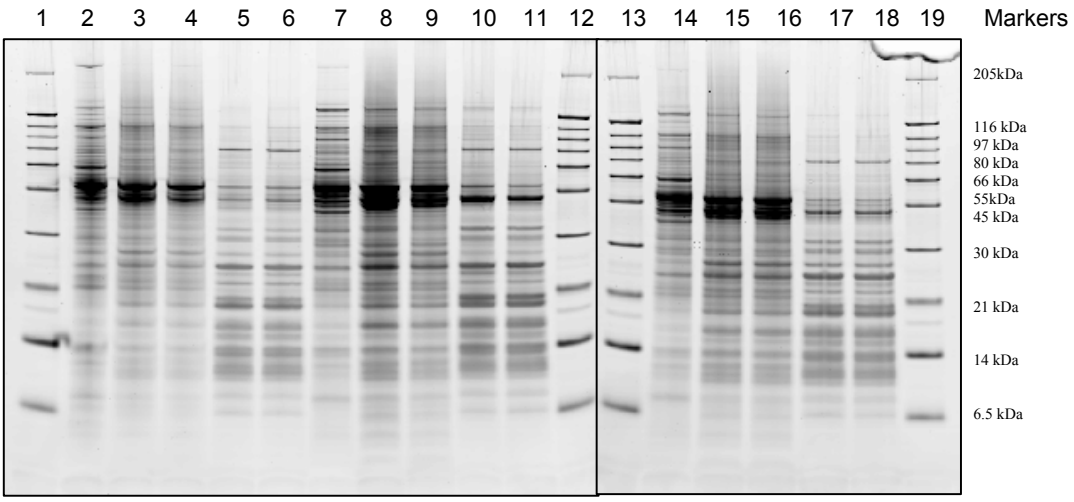


Figure 3.17 *– Comparison of the Sypro Ruby stained 1D profiles of both ProteoMiner treated fractions with crude protein extracts for each of the three strains.*
Lanes 1, 12, 13 and 19 contain a ladder of molecular markers. Lanes 2-6 correspond to the B-1 strain, with 2 being a B-1 crude protein extract, 3 + 4 E-bead treated B-1 extracts and 5 + 6 SE-bead treated B-1 extracts. Lanes 7-11 correspond to the A strain, with 7 being a crude A strain protein extract, 8 + 9 E-bead treated A extracts, and 10 + 11 SE-bead treated A extracts. Lanes 14-18 correspond to a second gel, showing the Tra5/5 strain extracts. Lane 14 contains a crude Tra5/5 strain protein extract, lanes 15 + 16 are E-bead treated Tra5/5 extracts and lanes 17 + 18 are SE-bead treated Tra5/5 extracts.

3.3.2 Differences in 2D profiles

The ProteoMiner treated and crude extracts were then compared by 2D gel electrophoresis (*Figure 3.18).*

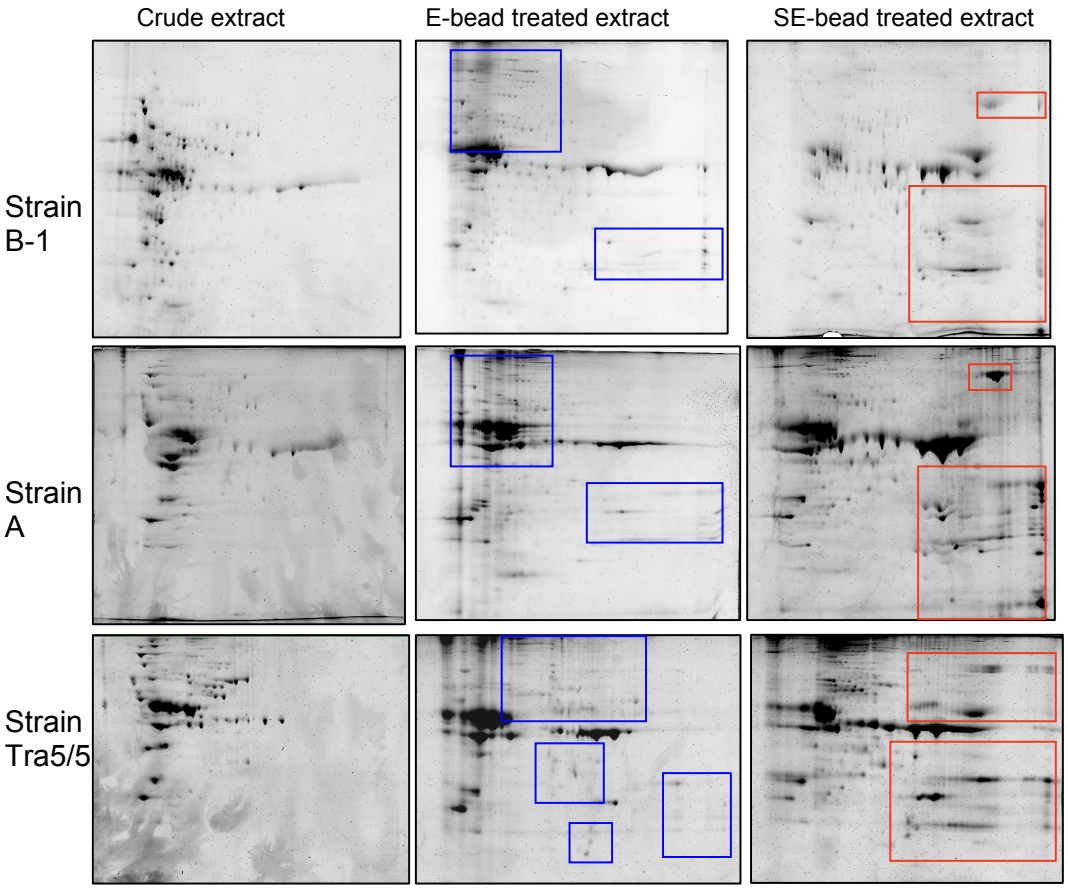


Figure 3.18 *– 2D gel electrophoresis of ProteoMiner treated and crude protein extracts for each of the three strains.*
For each of the three strains, crude, E-bead treated and SE-bead treated extracts were separated by 2D gel electrophoresis. The protein spots were visualised with SYPRO ruby protein stain. Areas of the gels showing additional proteins in treated samples are highlighted in blue boxes for E-bead treated fractions and red boxes for SE-bead treated fractions.

There are additional proteins present on the 2D gels of the ProteoMiner treated samples which are not visible on the gel profiles of crude extracts (*Figure 3.18*). Moreover, the additional proteins visualised in the E-bead and SE-bead treated fractions appear to be different, as was seen in the 1D gel profiles (*Figure 3.17*).

The protein extracts eluted from the SE-bead library with carboxylated terminal residues particularly show an increase of protein spots visible in the basic region of the gel compared with both the crude and E-bead treated fractions, as shown clearly by the positions of the red boxes on the SE-bead treated gels (*Figure 3.18).* Many of these are proteins of lower molecular mass, again confirming the differences in the 1D gel profiles (*Figure 3.17*). A number of these newly visualised proteins were picked from the gels for MALDI-TOF MS identification.

3.3.3 - Additional proteins identified in treated extracts by MALDI-TOF MS

For each strain, 48 additional spots were picked from the 2D gels of ProteoMiner treated fractions for MALDI-TOF mass spectrometry identification. These spots are shown in *Figures 3.19(i)-3.19(vi).* The majority of spots were recovered from the basic region of the gel, as these were considered of greater interest, having not previously been visualised on reference maps. For strain A, two SE-bead treated fraction profiles were chosen for spot picking, and for strains B-1 and Tra5/5, one E-bead treated fraction and one SE-bead treated fraction were used for spot picking.

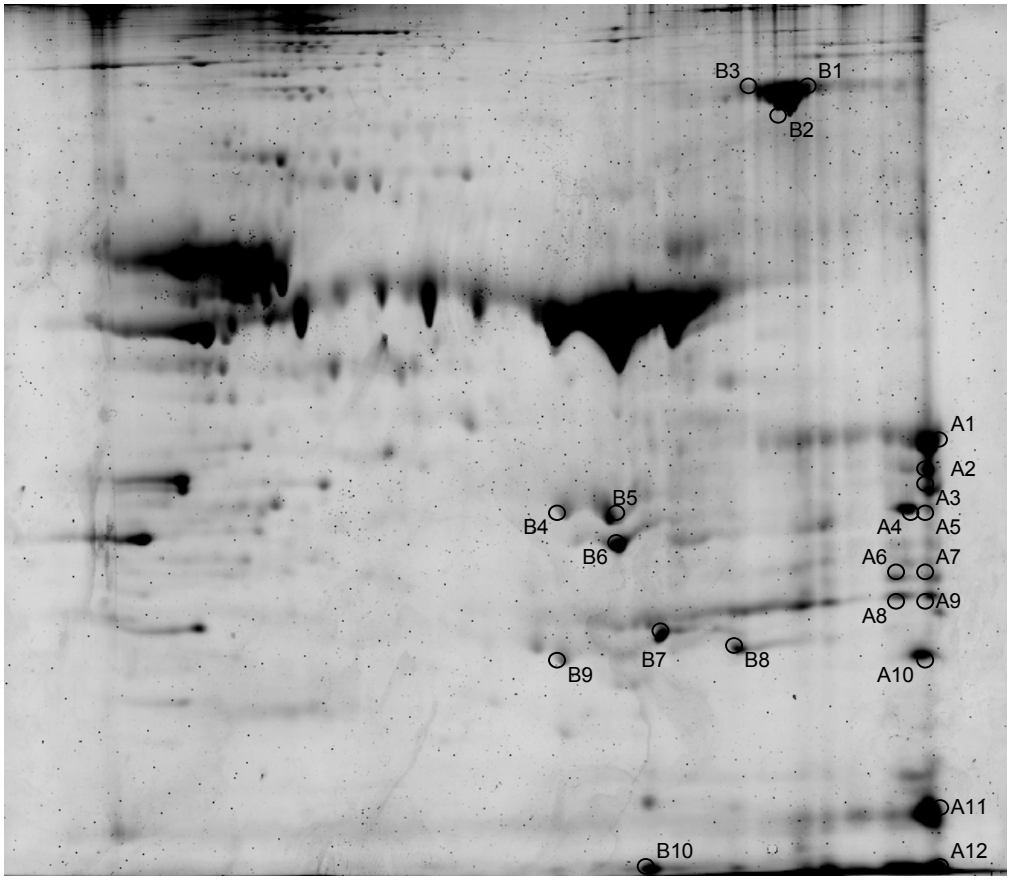


Figure 3.19(i) *– Additional protein spots for the strain A treated extracts excised from gel A, SE-bead fraction 1. Gels were post-stained with Brilliant Blue and the gel plugs were then digested for identification by MALDI-TOF MS. The additional proteins in the basic area of the gel were picked as these had not previously been visualised on 2D gels.*

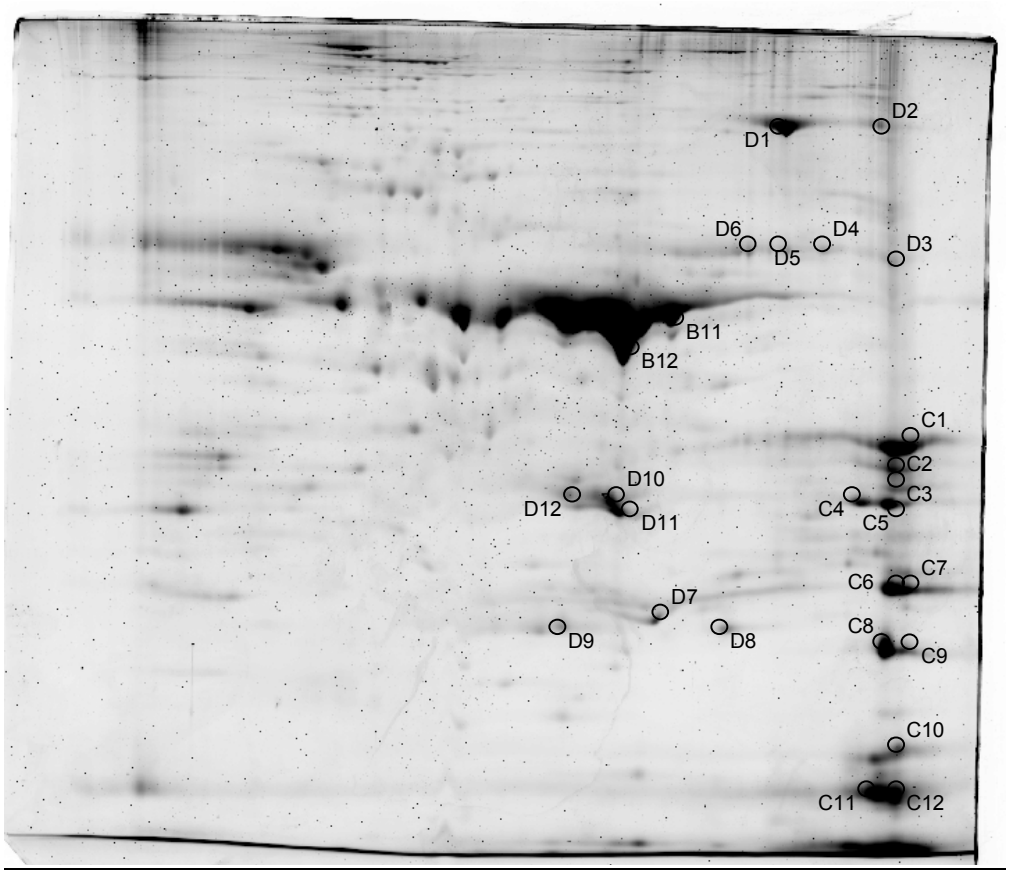


Figure 3.19(ii) *– Additional protein spots for the strain A treated extracts excised from gel B, SE-bead fraction 2. Gels were post-stained with Brilliant Blue and the gel plugs were then digested for identification by MALDI-TOF MS.*

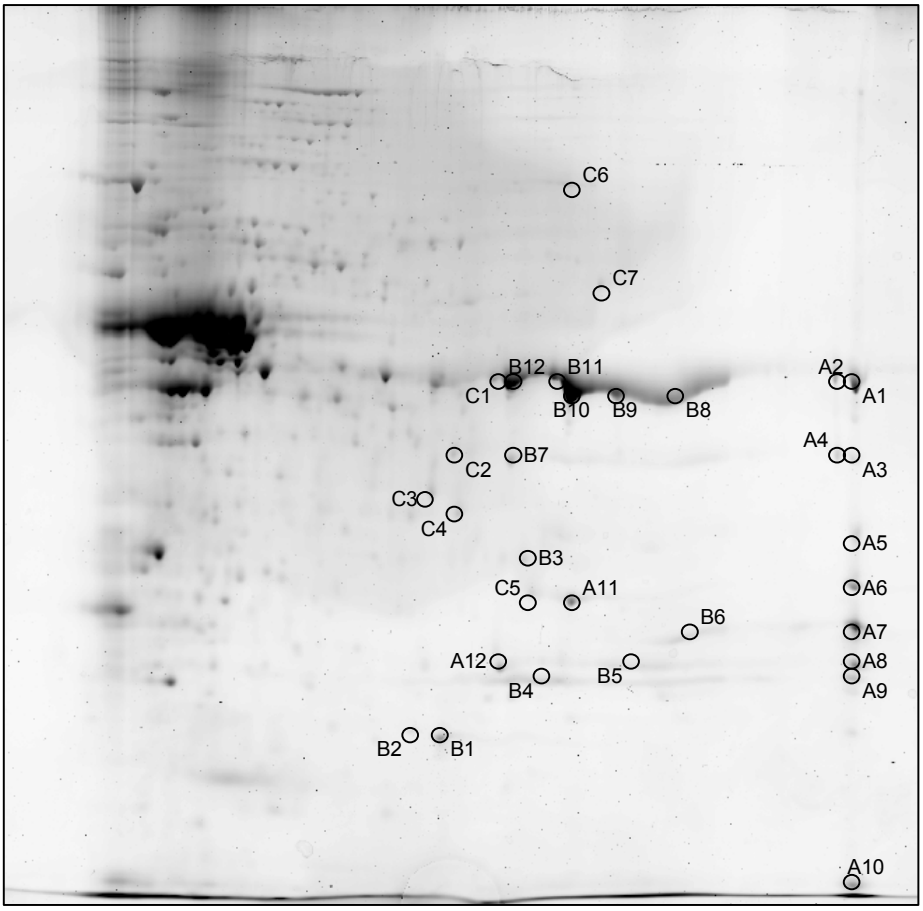


Figure 3.19(iii) *– Additional protein spots for the strain B-1 treated extracts excised from gel C, E-bead fraction. Gels were post-stained with Brilliant Blue and the gel plugs were then digested for identification by MALDI-TOF MS.*

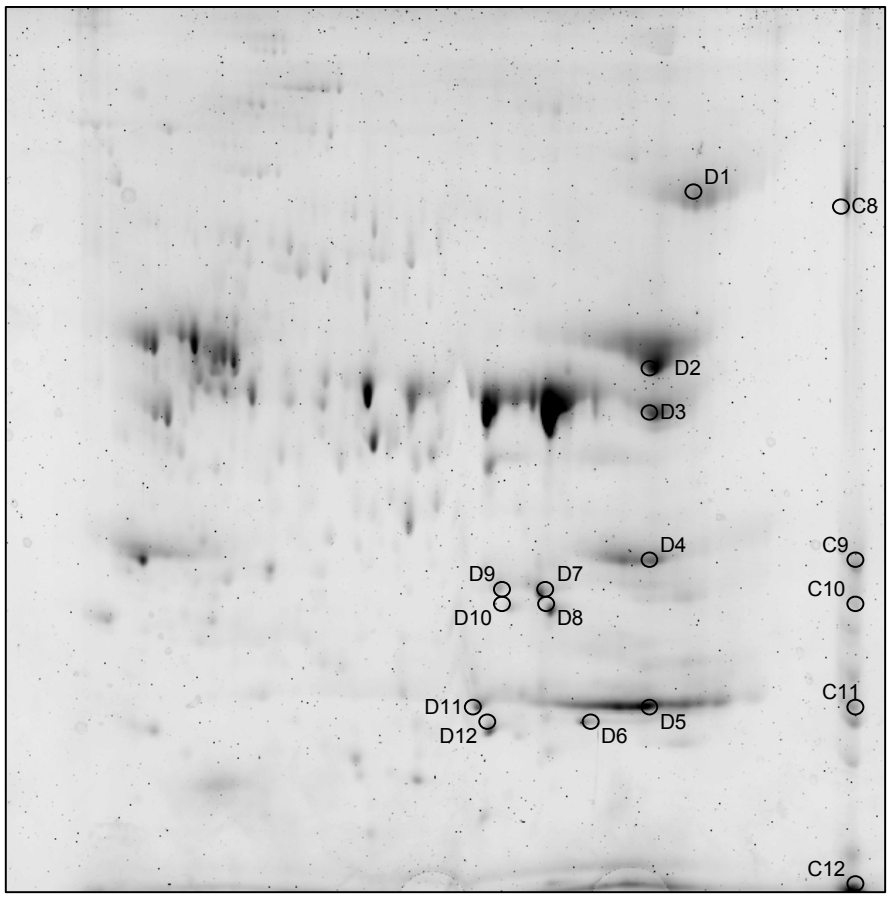


Figure 3.19(iv) *– Additional protein spots for the strain B-1 treated extracts excised from gel D, SE-bead fraction. Gels were post-stained with Brilliant Blue and the gel plugs were then digested for identification by MALDI-TOF MS.*

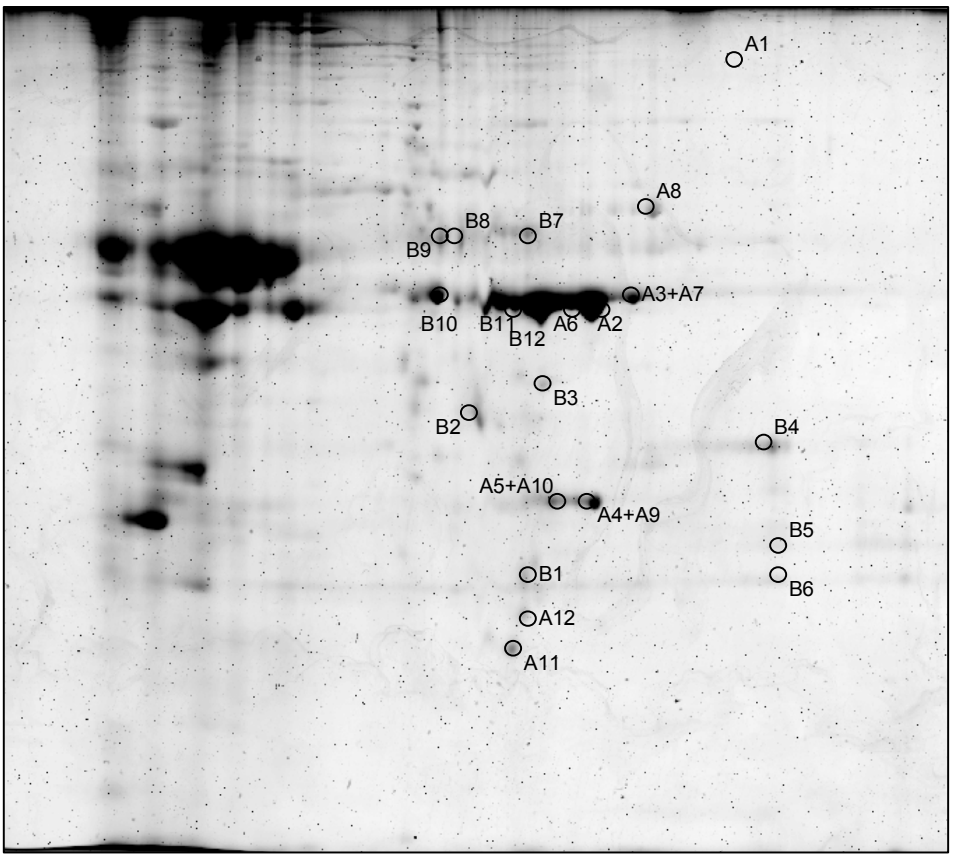


Figure 3.19(v) *– Additional protein spots for the strain Tra5/5 treated extracts excised from gel E, E-bead fraction. Gels were post-stained with Brilliant Blue and the gel plugs were then digested for identification by MALDI-TOF MS.*

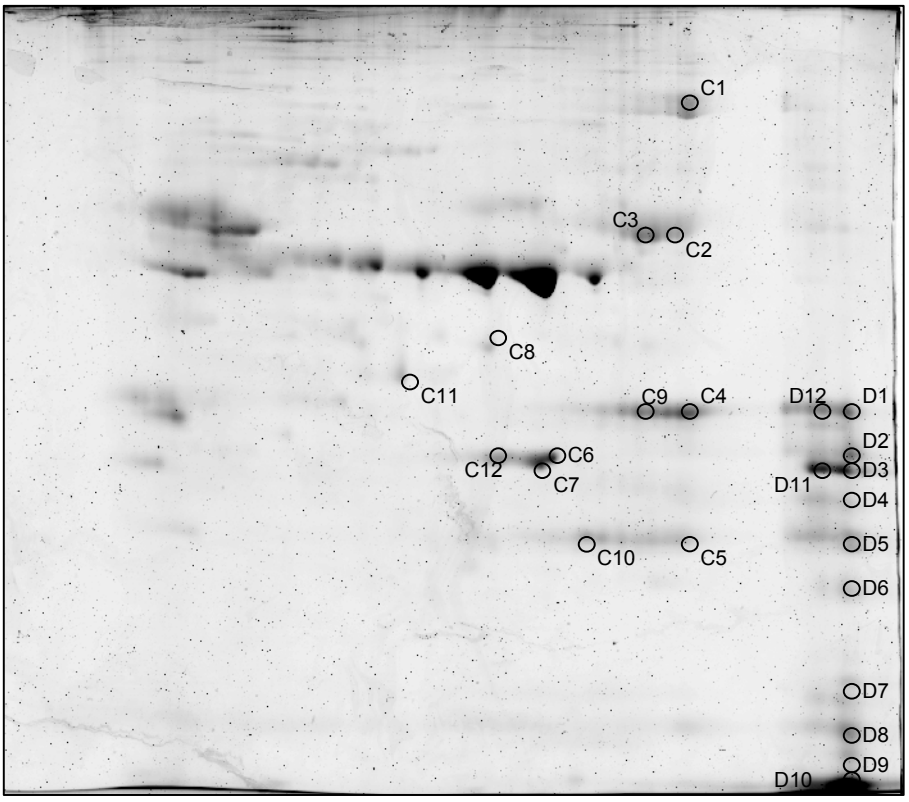


Figure 3.19(vi) *– Additional protein spots for the strain Tra5/5 treated extracts excised from gel F, SE-bead fraction. Gels were post-stained with Brilliant Blue and the gel plugs were then digested for identification by MALDI-TOF MS.*

For each strain, additional proteins were identified from the 2D profiles of the treated samples, some of which had not previously been identified on proteome reference maps of the strains.

For strain A, only 25 of the 48 additional spots picked from the 2D gels of ProteoMiner treated fractions were identified by MALDI-TOF MS. Thirteen different proteins were identified (*Table 3.12)*, only three of which had previously been identified on 2D reference maps, giving an additional ten

proteins detected by MALDI-TOF MS in the ProteoMiner treated fractions. Eight of these had previously been identified in the standard plate extracts using LC-MS/MS, but two proteins, a putative sugar-phosphate dehydrogenase and a terminase small subunit, have only been detected in the ProteoMiner treated strains.

For strain B-1, 43 of the 48 additional protein spots picked from the ProteoMiner 2D gels were identified by MALDI-TOF MS as 19 different proteins (*Table 3.13*). Six of these had been previously identified on the 2D reference maps, giving an additional 13 proteins detected by 2D reference mapping using ProteoMiner treatment. Of these 13 extra identifications, all had been previously identified in the standard plate extracts using LC-MS/MS.

For strain Tra5/5, 36 of the additional 48 spots picked from the ProteoMiner treated 2D gels were identified by MALDI-TOF MS as a total of 14 different proteins (*Table 3.14)*. Only four of these had previously been identified by 2D reference mapping, giving an extra ten proteins detected by MALDI-TOF MS in ProteoMiner treated samples. Nine of these had been previously identified using LC-MS/MS, but one protein, 50S ribosomal protein L19, had only been detected in the ProteoMiner treated samples.

Table 3.12 *– Strain A proteins identified by MALDI-TOF mass spectrometry from ProteoMiner treated protein extracts with MASCOT identification scores (A), theoretical mass (B) and pI (C), percentage peptide coverage (D) and number of matched peptides (E). Proteins shown in bold had not previously been identified on reference maps*

Strain A		Protein ID	number	A	B	C	D	E
A								
1	**A1 C1 C3**	**rpsC, 424475-425290 (Clockwise) 30S ribosomal protein S3**	**630_gi|115249083| emb|CAJ66894.1|**	**58**	**30401**	**10.33**	**31**	**8**
2	**A4 C4 C5**	**rplA, 405111-405809 (Clockwise) 50S ribosomal protein L1**	**630_gi|115249066| emb|CAJ66877.1|**	**87**	**24791**	**9.41**	**37**	**8**
3	B5 D10	spo0A, 1583111-1583935 (Clockwise) Stage 0 sporulation protein A homolog	630_gi|115250247| emb|CAJ68068.1|	82	30888	6.33	54	11
4	B6 D11	BASYS00470, 476328-477080 (Clockwise) Hypothetical Protein BASYS00470	630_gi|115249126| emb|CAJ66937.1|	62	27031	6.38	38	6
5	**B7 D7**	**tellurium resistance protein**	**630_gi|115250846| emb|CAJ68670.1|**	**89**	**21416**	**6.85**	**61**	**12**
6	**B8 D8**	**putative DNA-binding protein**	**630_gi|115249180| emb|CAJ66992.1|**	**61**	**23604**	**8.56**	**46**	**7**
7	B9	glyceraldehyde-3-phosphate dehydrogenase 2	630_gi|115252231| emb|CAJ70071.1|	59	36239	5.72	25	8
8	B11 B12	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405| emb|CAJ67220.1|	108	41469	5.88	50	17
9	**C6 C7 C10**	**50S ribosomal protein L4**	**630_gi|115249078| emb|CAJ66889.1|**	**89**	**23747**	**9.96**	**52**	**8**
10	**C8**	**50S ribosomal protein L5**	**630_gi|115249089| emb|CAJ66900.1**	**118**	**20424**	**9.47**	**67**	**16**
11	**C11**	**putative sugar-phosphate dehydrogenase**	**630_gi|115249504| emb|CAJ67319.1|**	**44**	**39053**	**5.89**	**27**	**6**
12	**D6**	**isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase**	**630_gi|115249401| emb|CAJ67216.1|**	**74**	**44491**	**5.18**	**35**	**10**
13	**B1 B3**	**1, 1756344-1756784 (Clockwise) Terminase small subunit**	**stb_BASYS01719_**	**60**	**16768**	**5.4**	**42**	**7**

Table 3.13 *– Strain B-1 proteins identified by MALDI-TOF mass spectrometry from ProteoMiner treated protein extracts with MASCOT identification scores (A), theoretical mass (B) and pI (C), percentage peptide coverage (D) and number of matched peptides (E). Proteins shown in bold had not previously been identified on reference maps.*

		Protein ID	number	A	B	C	D	E
B-1								
1	A1 A2 B8- C1	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405|emb|CAJ67220.1|	117	41469	5.88	49	16
2	A3 A4 B7 C2	fliC, 349752-348793 (CounterClockwise) Hypothetical Protein fliC	stb_BASYS00368	60	34339	7.82	41	8
3	**A5 C9 C10 D4**	**rpsC, 83439-84254 (Clockwise) 30S ribosomal protein S3**	**630_gi|115249083|emb|CAJ66894.1|**	**46**	**30401**	**10.33**	**26**	**7**
4	**A6**	**rplA, 64076-64774 (Clockwise) 50S ribosomal protein L1**	**630_gi|115249066|emb|CAJ66877.1|**	**62**	**24791**	**9.41**	**43**	**9**
5	**A7**	**rplC, 80301-80930 (Clockwise) 50S ribosomal protein L3**	**630_gi|115249077|emb|CAJ66888.1|**	**117**	**22370**	**9.95**	**56**	**15**
6	**A8 B5**	**rpsD, 93776-94351 (Clockwise) 30S ribosomal protein**	**630_gi|115249105|emb|CAJ66916.1|**	**68**	**21840**	**9.72**	**31**	**8**
7	**A9 B4 D5**	**rplD, 80930-81583 (Clockwise) Hypothetical Protein rplD**	**630_gi|115249078|emb|CAJ66889.1|**	**54**	**23747**	**9.96**	**43**	**7**
8	A11 C5 D8 D10	BASYS00279, 261373-260621 (CounterClockwise) Hypothetical Protein BASYS00279	630_gi|115249126|emb|CAJ66937.1|	78	27031	6.38	47	9
9	**A12 D2 D3 D11**	**isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase**	**630_gi|115249401|emb|CAJ67216.1|**	**77**	**44491**	**5.18**	**29**	**9**
10	**B1 B2**	**50S ribosomal protein L10**	**630_gi|115249067|emb|CAJ66878.1|**	**77**	**19581**	**6.2**	**50**	**7**
11	B3	rpsB, 2407846-2407133 (CounterClockwise) 30S ribosomal protein	630_gi|115251194|emb|CAJ69025.1|	62	26940	6.19	44	9
12	C3	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079|emb|CAJ67899.1|	70	30887	5.8	28	8
13	C4	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249404|emb|CAJ67219.1|	40	42966	5.27	15	5
14		**manganese-dependent inorganic pyrophosphatase**	**630_gi|115249342|emb|CAJ67155.1|**	**57**	**59668**	**4.89**	**34**	**11**

15		putative GTP-binding protein	630_gi|115251683|emb|CAJ69518.1|	89	49681	6.51	28	13
16	C12	50S ribosomal protein L15	630_gi|115249096|emb|CAJ66907.1|	58	16142	10.44	52	9
17	D1	cell surface protein	630_gi|115251764|emb|CAJ69599.1|	60	72030	8.89	33	15
18	D6	tellurium resistance protein	630_gi|115250846|emb|CAJ68670.1|	114	21416	6.85	59	12
19	D7 D9	spo0A, 1360119-1360943 (Clockwise) Stage 0 sporulation protein A homolog	630_gi|115250247|emb|CAJ68068.1|	166	30888	6.33	83	18

Table 3.14 *– Strain Tra5/5 proteins identified by MALDI-TOF mass spectrometry from ProteoMiner treated protein extracts with MASCOT identification scores (A), theoretical mass (B) and pI (C), percentage peptide coverage (D) and number of matched peptides (E). Proteins shown in Bold had not previously been identified on reference maps.*

Strain Tra 5/5		Protein ID	number	A	B	C	D	E
1	A2 A3 A6 A7 A12 B11 B12	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405|emb|CAJ67220.1|	129	41469	5.88	49	18
2	A4 A5 A9 A10 C6	putative subunit of oxidoreductase	630_gi|115249126|emb|CAJ66937.1	123	27031	6.38	56	11
3	**A8**	**putative GTP-binding protein**	**630_gi|115251683|emb|CAJ69518.1**	**225**	**49681**	**6.51**	**58**	**25**
4	**A11**	**50S ribosomal protein L10**	**630_gi|115249067|emb|CAJ66878.1**	**137**	**19581**	**6.2**	**80**	**15**
5	**B1 C10**	**50S ribosomal protein L4**	**630_gi|115249078|emb|CAJ66889.1|**	**47**	**23747**	**9.96**	**47**	**7**
6	B2	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079|emb|CAJ67899.1|	102	30887	5.8	39	11
7	**B4 C4 D1 D2 D4**	**30S ribosomal protein S3**	**630_gi|115249083|emb|CAJ66894.1|**	**153**	**30401**	**10.33**	**59**	**17**
8	**B6**	**30S ribosomal protein S4**	**630_gi|115249105|emb|CAJ66916.1|**	**86**	**21840**	**9.72**	**45**	**10**
9	B10	butyryl-CoA dehydrogenase	630_gi|115250075|emb|CAJ67895.1|	146	41390	5.71	55	17
10	**C1**	**cell surface protein**	**630_gi|115251764|emb|CAJ69599.1|**	**109**	**72030**	**8.89**	**30**	**19**
11	**C2 C3**	**isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase**	**630_gi|115249401|emb|CAJ67216.1|**	**90**	**44491**	**5.18**	**33**	**12**
12	C7	spo0A, 1301769-1302593 (Clockwise) Stage 0 sporulation protein A homolog	630_gi|115250247|emb|CAJ68068.1|	84	30888	6.33	55	12
13	**D3 D11**	**50S ribosomal protein L1**	**630_gi|115249066|emb|CAJ66877.1|**	**129**	**24792**	**9.32**	**61**	**15**

14	**D10**	**rplS, 1334033-1334383 (Clockwise) 50S ribosomal protein L19**	**630_gi|115250291|emb|CAJ68113.1**	**84**	**13287**	**10.84**	**60**	**12**

Few additional surface or virulence proteins were detected in the 2D profiles of treated strains, but 2D GE followed by MALDI-TOF MS identified one sporulation potein, Spo0A in the treated extracts of all three strains, whereas previously it had not been detected in strain B-1, and an additional cell surface protein (gi|115251764) in strains B-1 and Tra5/5 but not A. These had both been identified in the standard plate extracts by LC-MS/MS.

3.3.4 Additional proteins identified in treated extracts by LC-MS/MS

In order to get a more in-depth picture of the changes occurring due to the ProteoMiner treatment, the 1D gel profiles of crude and treated extracts (*Figure 3.17*) were cut into 12 bands, and each band subject to in-gel digestion before sizing on the LTQ-orbitrap quadrapole mass spectrometer for a more sensitive identification of proteins present in each fraction. The raw data files were searched against a redundant database of strains A, B-1, Tra5/5 and the reference strain 630 using MASCOT. Protein identifications were exported from scaffold and all those identification which were false were deleted. As with other mascot identifications, all protein identifications based on 1 peptide only were verified by searching the corresponding peptide against the translated genome of the corresponding strain. If the peptide was present in the relative strain, the identification was considered valid. If the peptide was not present, the identification was discarded. The identified proteins are shown in *Appendices A5, A6 and A7.*

For strain A, a total of 664 proteins were identified by Scaffold, 19 of which were discarded as false identifications. After redundancy was removed, 643 identifications were confirmed as shown in *Figure 3.20* below.

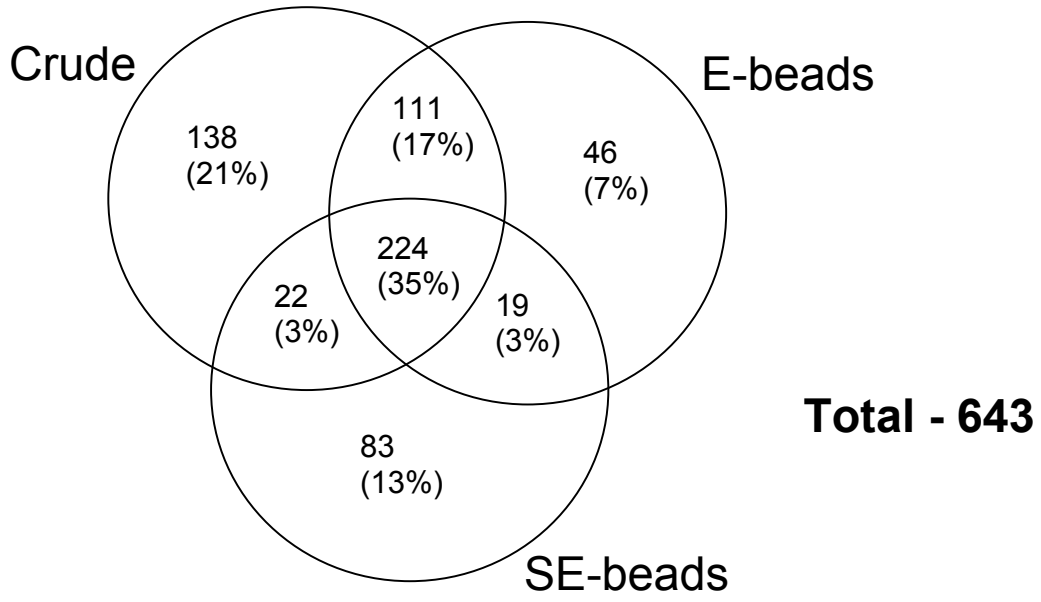


Figure 3.20 *– A Venn diagram to show the breakdown of proteins identified in the strain A crude and ProteoMiner treated extracts by LC-MS/MS. If a protein was present in one replicate, it was considered to be present in the sample. Identifications, theoretical mass and number of unique peptides are shown in Appendix A5.*

For strain B-1, a total of 558 proteins were identified by Scaffold, 54 of which were discarded as false identifications. After removing redundancy, 503 identifications were confirmed as shown in *Figure 3.21* below.

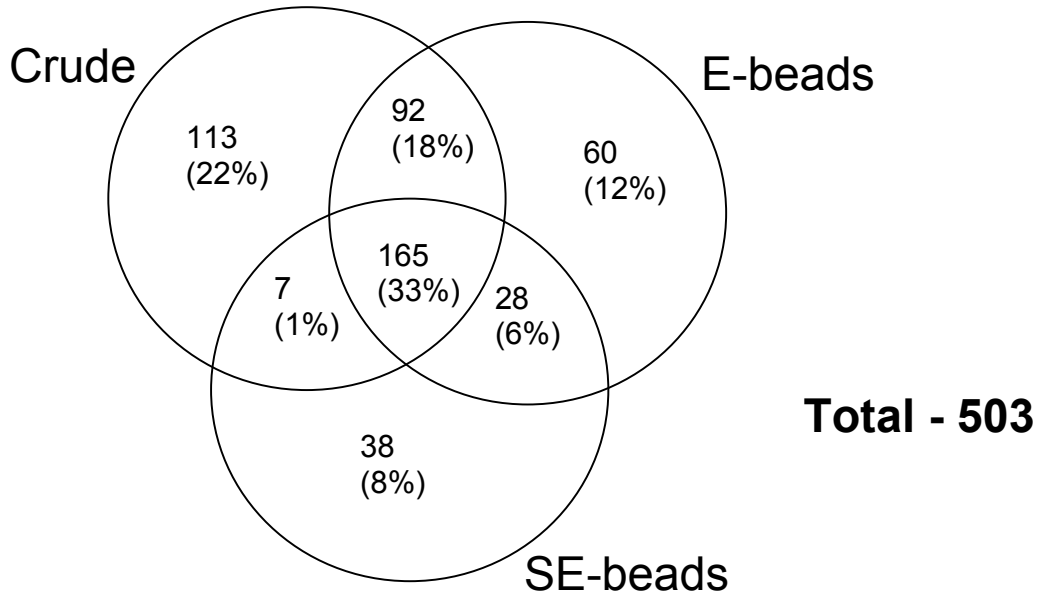


Figure 3.21 *– A Venn diagram to show the breakdown of proteins identified in the strain B-1 crude and ProteoMiner treated extracts by LC-MS/MS. If a protein was present in one replicate, it was considered to be present in the sample. Identifications, theoretical mass and number of unique peptides are shown in Appendix A6.*

For strain Tra5/5, a total of 622 proteins were identified by Scaffold, 22 of which were discarded as false identifications. After redundancy was removed, 599 identifications were confirmed as shown in *Figure 3.22* below.

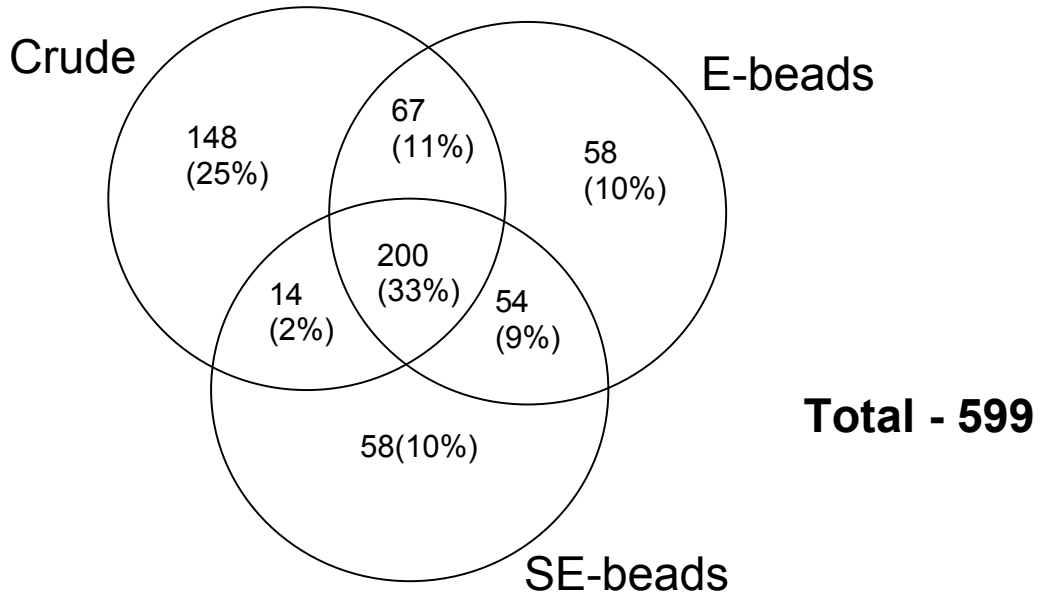


Figure 3.22 *– A Venn diagram to show the breakdown of proteins identified in the strain Tra5/5 crude and ProteoMiner treated extracts by LC-MS/MS. If a protein was present in one replicate, it was considered to be present in the sample. Identifications, theoretical mass and number of unique peptides are shown in Appendix A7.*

3.3.4.1 - Comparison of LC-MS/MS data for broth and standard plate extracts

The protein samples treated with the ProteoMiner bead libraries were extracted from broth cultures whereas the standard protein extracts used for protein mapping and 1D LC-MS/MS were from plate culture. Growth conditions have a large effect on protein expression, so the proteins detected in the crude broth extracts by LC-MS/MS were compared to the proteins detected in plate extracts by LC-MS/MS.

In strain A, 360 proteins were identified from the plate extracts which were not detected in the broth extracts used for ProteoMiner treatment, and in strains B-1 and Tra5/5 it was 428 and 235 proteins respectively. This constituted 57%, 64% and 55% respectively of the total proteins detected in the standard plate extracts

The additional proteins detected in the broth extract used for ProteoMiner treatment but not in the standard plate extracts were then looked at in more detail and are shown in the *Table 3.15* below

Table 3.15 *– Proteins detected in crude broth samples but not the standard plate extracts.*

	Additional proteins detected in broth extracts but not standard plate extracts		
	Strain A	**Strain B-1**	**Strain Tra5/5**
All samples	81	48	97
E and Crude	54	45	37
E only	36	42	50
E and SE	16	14	37
SE only	74	30	49
SE and Crude	11	3	11
Crude only	78	61	117
Total (% of total proteins detected in broth extracts)	350 (54%)	243 (48%)	398 (66%)

For strain A, 350 proteins were identified in the broth extracts which were not detected in the standard plate extracts, of which 126 were found only in ProteoMiner treated samples. For strain B-1, 243 proteins were identified in the broth extracts which were not detected in the standard extracts, of which 86 were seen only in ProteoMiner treated fractions. For strain Tra5/5, 398 proteins were identified in the broth extracts which were not detected in the standard plate extracts, 136 of which were found only in the ProteoMiner treated samples.

3.3.4.2 - Affects of ProteoMiner treatment on the broth extracts

The protein identifications from the crude broth extracts and E-bead and SE-bead treated samples were investigated in more detail to determine how many additional proteins were detected after ProteoMiner treatment

A total of 148 additional proteins were present in strain A after ProteoMiner treatment that were not detected in the crude protein extract. These included 46 identified only in the E-bead treated fraction, 83 identified only in the SE-bead

treated fraction, and 19 identified in both treated fractions. Of these, eight were surface or virulence proteins, including one sporulation protein (Spo0B-assosiated GTP-binding protein-gi|115250196) found in both E-bead and SE-bead treated fractions and not detected in the standard plate extracts for this strain. Two sporulation proteins (gi|115251680, gi|115250981), four cell surface proteins (gi|115251032, gi|115251842, gi|115250795, gi|115250166) and a flagellar protein (gi|115249243) were found in the SE-bead treated fraction only. The flagellar protein and one of the surface proteins (gi|115251032) were not detected at all in the standard plate extracts. There were also three ABC transporter proteins found only in ProteoMiner treated samples. However, 138 proteins present in the crude extract were lost after treatment.

A total of 127 additional B-1 proteins were present after ProteoMiner treatment that were not detected in the crude protein extract. These included 60 identified only in the E-bead treated fraction, 38 identified only in the SE-bead treated fraction, and 29 identified in both treated fractions. The flagellin subunit was found in both the E-bead and SE-bead treated fractions, and two surface proteins (gi|115251847 and gi|115251846) were found only in the SE-bead treated fraction. One of these (gi|115251847) had not previously been identified in strain B-1 in the standard plate extracts. However, 113 (22%) of the proteins identified in the crude extract were lost after ProteoMiner treatment.

In strain Tra5/5, a total of 171 additional proteins were present in the ProteoMiner treated samples which were not detected in the crude broth extracts. These included 58 proteins identified only in the E-bead treated fraction, 58 identified only in the SE-bead treated fractions, and 54 identified in

both ProteoMiner treated fractions. A sporulation protein (gi|115252560) found only in the E-bead treated fraction and not identified in the standard plate extracts for this strain, a cell surface protein (gi|115251842) found in both treated fractions and detected in the standard plate extracts, and two sporulation related proteins (gi|115250981, gi|115250196) and a putative flagellar protein (gi|115249264) found in only the SE-bead treated fraction, only one of which (gi|115250196) had been seen in the standard plate extracts for this strain. The putative flagellar protein (gi|115249264) was not detected in any of the standard plate extracts. However, 148 (25%) of the proteins identified in the crude extract were lost after ProteoMiner treatment including the pilin protein detected in strain Tra5/5 in the crude broth samples but not in the standard plate extracts or the ProteoMiner treated samples. Pilin was not seen in any samples of the other two strains.

3.4 – Western Blotting

3.4.1 – Horse Serum

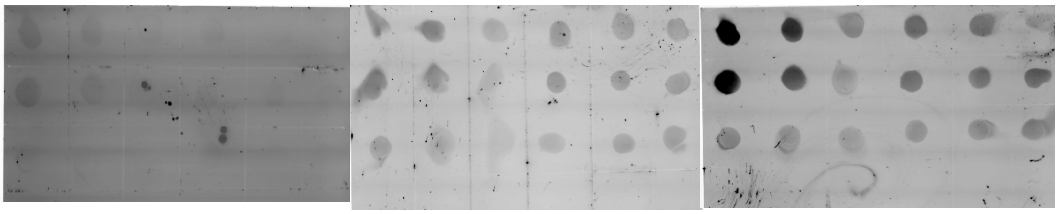
In order to optimise the Western blotting procedures, horse serum was used as a protein sample. A rabbit anti-horse primary antibody was used (SIGMA) in conjunction with a goat polyclonal to rabbit IgG secondary antibody. CyDye labelled secondary antibodies were used rather than chemiluminescent methodologies, because of their compatibility with the Ettan Dalt imager.

Dot blots with serial dilutions of primary antibody (rabbit anti-horse) were incubated with Cy3 labelled secondary antibody (goat polyclonal to rabbit IgG)

at 1:20,000, 1:10,000 and 1:5000 and shown in *Figure 3.23* below. There was a range of concentrations of both primary and secondary antibodies that allowed easy visualisation of immunoreactive proteins.

1/50	1/100	1/500	1/1000	1/2000	1/4000
1/50	1/100	1/500	1/1000	1/2000	1/4000
1/5000	1/5000	1/8000	1/8000	1/10000	1/10000

i) Primary antibody (rabbit anti-horse) dilutions used in dot blots



1:20,000 1:10,000 1:5000

ii) Dot blots of the primary antibody dilutions (i) with a range of concentrations of Secondary Antibody (goat polyclonal to rabbit IgG)

***Figure 3.23** – Dot Bots of varying concentrations of primary and secondary antibodies used for Western blotting of horse serum proteins. The different primary antibodies used are shown in (i), and blots of these dilutions incubated at three different concentrations of secondary antibody are shown in (ii).*

In order to optimise blotting conditions, serial dilutions of the horse serum (5µl) were run on duplicate 12% NuPage gels alongside a molecular ladder (Protein Molecular Weight Standards (Molecular Probes P-6649)). These proteins were detected with the rabbit anti-horse primary antibody, and the Cy3 labelled anti-

rabbit secondary antibody. Different staining and labelling methods were trialled, and a protocol optimised.

Where proteins were unlabelled, two gels were run in parallel. One gel was fixed and stained with Sypro Ruby, and the proteins from the second gel were electro-transferred to nitrocellulose membrane using a current of 45mA for 1 hour. Transfer was confirmed by staining the membrane with Ponceau S solution or SYPRO Ruby blot stain before probing with primary and secondary antibodies.

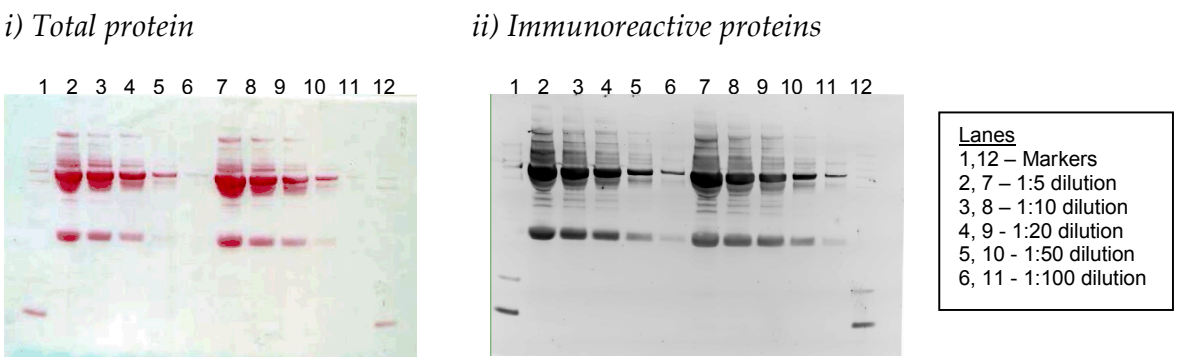


Figure 3.24 *–1D Western blots of horse serum dilutions probed with rabbit anti-horse. Protein sample - 5µl serial dilution of horse serum; Primary antibody – Rabbit anti-horse, 1:2000 dilution, incubated for 2hours; Secondary antibody – Cy3 labelled goat anti-rabbit, 1:10,000 dilution, overnight incubation. (i) Total protein stained with Ponceau S post-blotting; (ii) Cy5-labelled immunoreactive proteins.*

Firstly, unlabelled serial dilutions of horse serum proteins were separated, blotted to nitrocellulose membrane and probed with primary and secondary antibodies (*Figure 3.24*). As expected, all horse serum proteins were recognised

by the antibodies, but there was also some non-specific binding to the marker proteins (*Figure 3.24)*. The smallest protein in the molecular markers was Aprotinin from horse muscle, so would be expected to be recognised by the anti-horse primary antibody. However, this was not the only marker protein to be recognised by labelled antibodies.

In order to determine the extent of the non-specific binding, *Clostridium difficile* protein extracts were used as a negative control (*Figure 3.25)*, as they should not be recognised by the rabbit anti-horse primary antibody. After probing, all the horse serum proteins had been recognised by the antibodies (*Figure 3.25*), and there was no non-specific binding to the *Clostridium difficile* extracts.

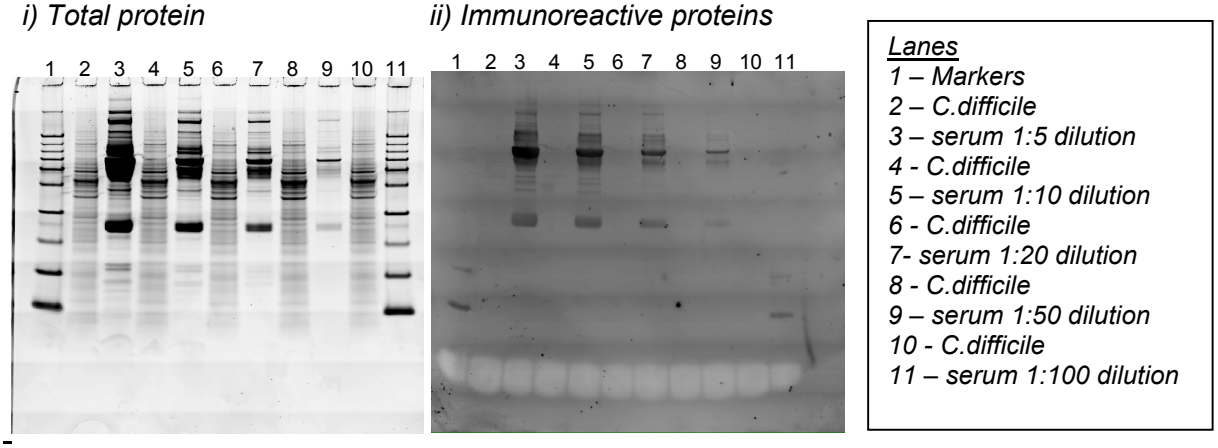


Figure 3.25 *– 1D Western blots of horse serum and C. difficile protein extracts, probed with an anti-rabbit primary antibody. Protein sample - 5µl serial dilution of horse serum, or 5µg Clostridium difficile protein extract; Primary antibody – Rabbit anti-horse, 1:5000 dilution, incubated for 2hours; Secondary antibody – Cy3 labelled goat anti-rabbit, 1:20,000 dilution, 2hour incubation. (i) Total protein from second, SYPRO strained gel; (ii) Cy5 labelled secondary antibodies from blot.*

In order to visualise total protein and immunoreactive proteins together, the membrane was stained with Sypro Ruby before probing. Cy3 has the same emission wavelength as Sypro Ruby, so Cy5 labelled secondary antibodies were used instead. Using the Ettan DALT, two wavelengths can be scanned at the same time, so immunoreactive proteins can be matched to total protein quickly and easily in the same image (*Figure 3.26*). Total protein stained with Sypro is visible in green, and the Cy5 labelled secondary antibodies are visible in red. Where the two overlap, the protein bands appear yellow. When proteins are bound strongly by the antibodies, they appear red, although on separate images, the corresponding green bands can still be visualised (*Figure 3.26)*.

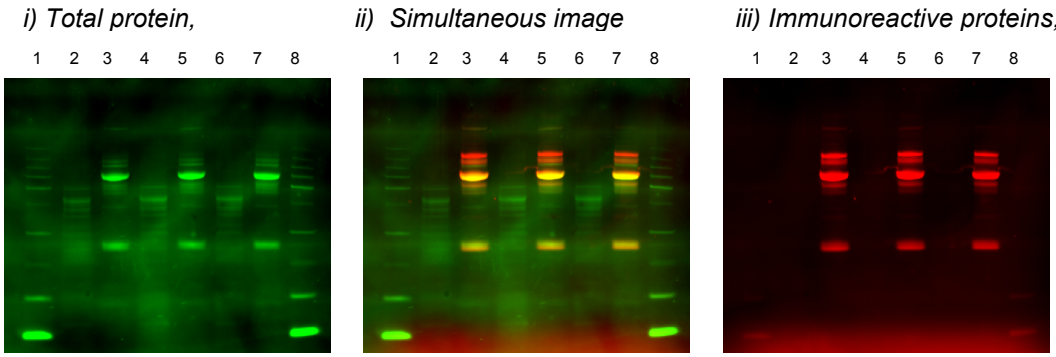


Figure 3.26 *– Blots of Sypro Ruby stained horse serum and Clostridium difficile extracts probed with rabbit anti-horse.*
Lanes – 1+8 protein markers, 2+4+6 – 5µg Clostridium difficile extract, 3+5+7 - 5µg horse serum protein. Primary antibody – Rabbit anti-horse, 1:5000 dilution, incubated for 2hours; Secondary antibody – Cy5 labelled goat anti-rabbit, 1:20,000 dilution, 2hour incubation. (i) Total protein blot strained with SYPRO ruby blot stain; (ii) SYPRO channel and CY5 channel imaged together; (iii) Immunoreactive proteins, Cy5 labelled secondary antibody.

In order to improve the sensitivity of the staining of the total protein samples, both the horse serum and the *C. difficile* proteins extracts were labelled with Cy2 (*Figure 3.27*). This allows the total protein profile to be visualised in the gel before transfer, in the membrane after transfer, and together with the secondary antibody in the same image after probing, as Cy2 and Cy5 have different emission wavelengths. This saves time and resources as there is no need for duplicate gels, and, due to the high sensitivity of CyDyes, allows greater visualisation of the total protein.

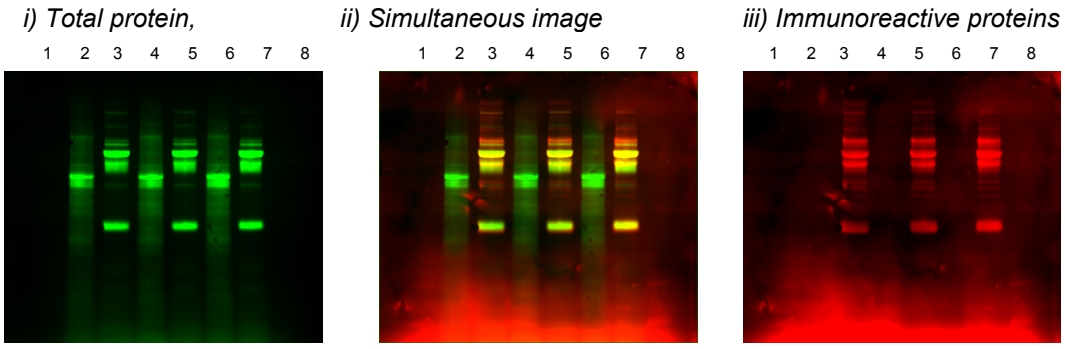


Figure 3.27 *– 1D Western blots of CyDye labelled horse serum and Clostridium difficile extracts probed with rabbit anti-horse. Lanes – 1+8 protein markers, 2+4+6 – 5µg Cy2 labelled Clostridium difficile extract, 3+5+6- 5µg Cy2 labelled horse serum protein; Primary antibody – Rabbit anti-horse, 1:5000 dilution, incubated for 2hours; Secondary antibody – Cy5 labelled goat anti-rabbit, 1:20,000 dilution, 2hour incubation. (i) Total protein extract labelled with Cy2 before electrophoresis; (ii) Cy2 and Cy5 channel imaged together; (iii) Immunoreactive proteins, Cy5 labelled secondary antibody.*

3.4.2 – Rabbit Serum

Rabbit serum was obtained from Professor Neil Fairweather at Imperial College London. Antibodies against *Clostridium difficile* strain 630 were raised by injecting rabbits with a heat killed preparation of whole bacterial cells.

The rabbit serum had previously been used at a 1:20,000 dilution (personal communication) and dot bots were used to confirm this was the optimal serum concentration (*Figure 3.28*). The dot blots indicated that a higher concentration would be optimal, but good results for 1D Western blots were obtained with both a 1:10,000 and 1:20,000 dilution as discussed below.

Protein extracts of the three different *Clostridium difficile* cells were separated by 1D gel electrophoresis, and the rabbit serum was used as the primary antibody at the 1:20,000 dilution suggested. The anti- rabbit IgG secondary antibody was used at the same concentration as above (1:20,000). Protein transfer was confirmed by Sypro straining, and immunoreactive proteins visualised by the labelled secondary antibodies. A number of immunoreactive proteins were recognised by the rabbit serum, and there was a clear difference in the immunoreactive proteins in strain B-1 compared to strain A and Tra5/5 (*Figure 3.28*).

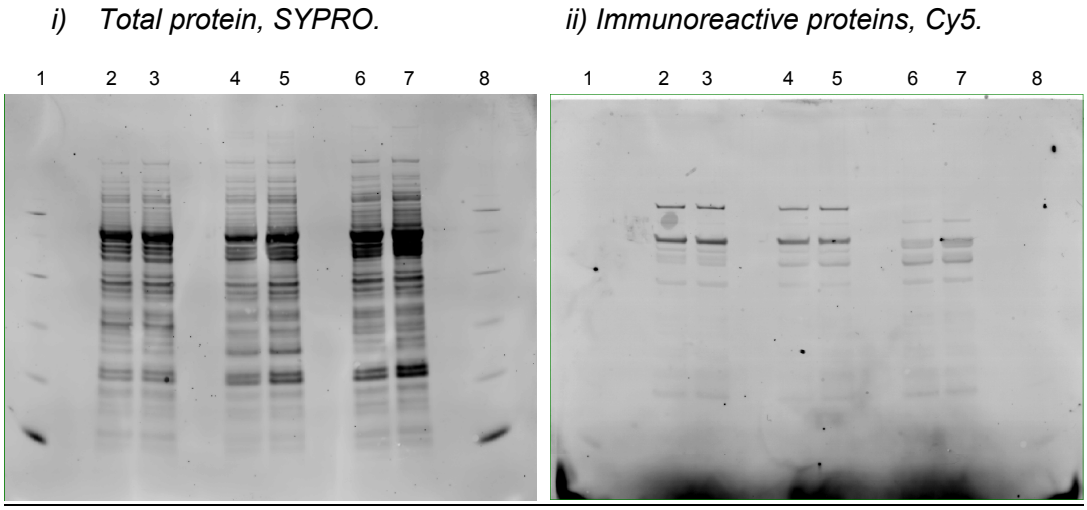


Figure 3.28 *– 1D Western blots of unlabelled C. difficile extracts probed with rabbit serum. Lanes - 1+8 molecular ladder, 2+3 - 5µg Clostridium difficile strain Tra5/5 extract, 4+5 - 5µg Clostridium difficile strain A extract, 6+7– 5µg Clostridium difficile B-1 extract. Primary antibody - Rabbit serum, 1:20,000 dilution, incubated for 2hours; Secondary antibody – Cy5 labelled goat anti-rabbit, 1:20,000 dilution, 2hour incubation. (i) Total protein strained with SYPRO; (ii) Immunoreactive proteins recognised by Cy5 secondary antibodies.*

The *C. difficile* proteins extracts were labelled as above with Cy2 before electrophoresis to allow total proteins and immunoreactive proteins to be visualised together (*Figure 3.29*).

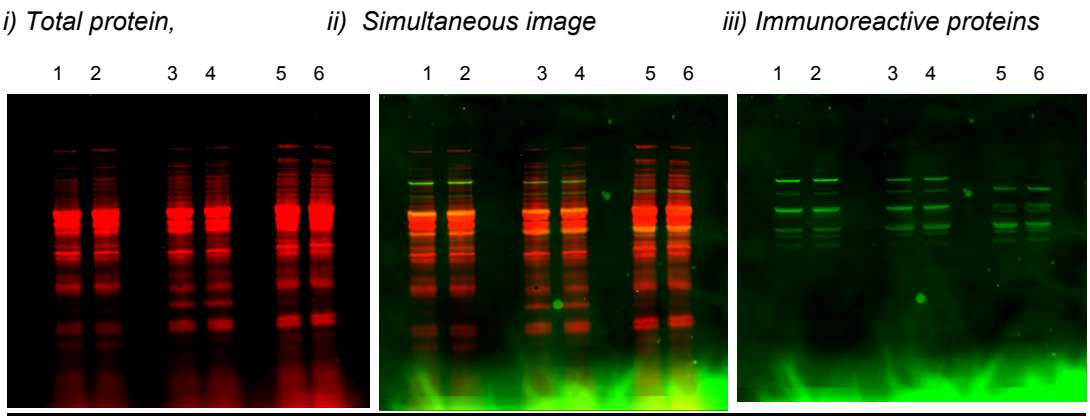


Figure 3.29 *– 1D Western blots of Cy2 labelled Clostridium difficile protein extracts probed with rabbit serum. Lanes – 1+2 - 5µg Cy2 labelled Clostridium difficile strain Tra5/5 extract, 3+4 - 5µg Cy2 labelled Clostridium difficile strain A extract 5+6 - 5µg Cy2 labelled Clostridium difficile strain B-1 extract. Primary antibody – Rabbit serum, 1:20,000 dilution, incubated for 2hours; Secondary antibody – Cy5 labelled goat anti-rabbit, 1:20,000 dilution, 2hour incubation. (i) Total protein extract labelled with Cy2 before electrophoresis; (ii) Cy2 and Cy5 channel imaged together; (iii) immunoreactive proteins, Cy5 labelled secondary antibody.*

The best 1D gel result was obtained using a higher concentration of primary antibody in 10 mM EDTA (*Figure 3.30)*.

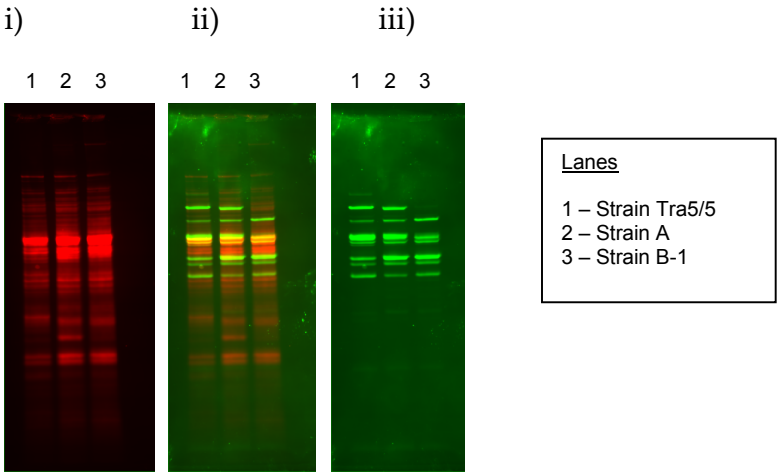


Figure 3.30 *– Cy2 labelled protein extracts probed with rabbit serum using EDTA Protein sample - 5µg Cy2 labelled Clostridium difficile protein extract from three different strains. Membranes were blocked with 1% BSA overnight; Primary antibody – Rabbit serum, 1:10,000 dilution in 0.01M PBS Tween, 0.5% BSA, 10mM EDTA, incubated for 6 hours. Membranes were washed 3 x 5 minutes in 0.01M PBS Tween; Secondary antibody – Cy5 labelled goat anti-rabbit, 1:20,000 dilution in 0.01 M PBS Tween , 0.5% EDTA, 2 hour incubation. (i) Total protein extract labelled with Cy2 before electrophoresis; (ii) Cy2 and Cy5 channel imaged together; (iii) Immunoreactive proteins, Cy5 labelled secondary antibody*

Because of the complexity of the whole cell protein extracts, it is difficult to tell which protein band matches the immunoreactive proteins, so identification is problematic, but visualising both the total protein extract and the immunoreactive proteins together allows some identification of those proteins recognised by the rabbit immune system by determining approximate molecular weights from the molecular markers (*Figure 3.31*). Although the markers are not Cy5 labelled, so cannot be visualised with the CyDye labelled proteins, the profiles of the *C. difficile* extracts can be matched to

those run alongside markers to give approximate sizes, as shown below (*Figure 3.31*).

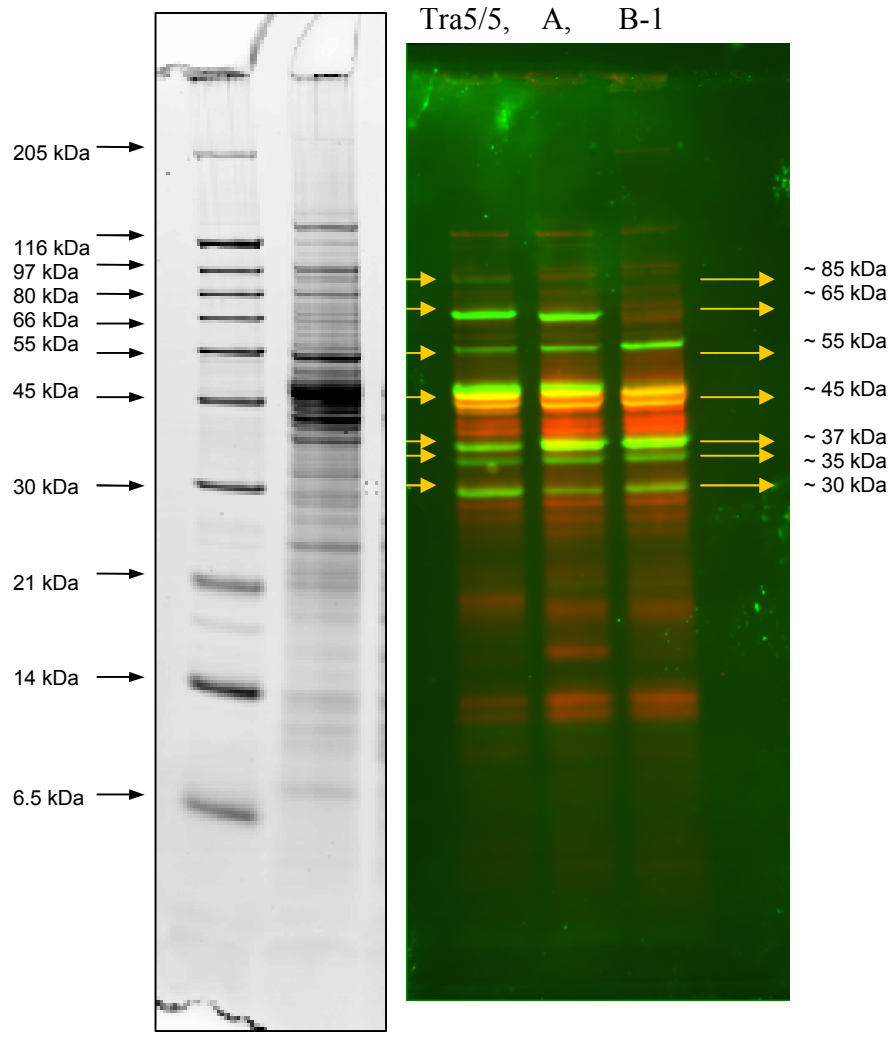


Figure 3.31 *– 1D Western blotting profiles of each strain matched to the molecular markers to determine approximate molecular weights of the immunoreactive proteins detected. The immunorective protein profile shown in Figure 3.30 (ii) was matched to a Sypro stained strain A whole cell protein extract next to a molecular marker ladder (i). This allows the Cy2 stained whole cell protein extracts in blot (ii) (shown in red) to be matched to the extract in blot (i) to give approximate sizes of the immunoreactive bands (Cy5 labelled secondary antibody shown in green). Immunoreactive bands are highlighted by yellow arrows with corresponding approximate sizes.*

The high molecular weight S-layer protein was expected to be ~ 45 kDa, and the low molecular weight S-layer protein was expected to be ~35 kDa (Calabi *et al*, 2001). Immunoreactive protein bands corresponding to high abundance proteins can be seen at both these approximate positions, corresponding to several proteins of ~45kDa and two strongly reactive proteins of approximately 35 kDa. In addition, immunoreactive bands of approximately 85 kDa, 65 kDa, 55 kDa, and 30 kDa were present. The band at ~85kDa was present only in strain Tra5/5, while the band at ~65 kDa was present in strains Tra5/5 and A but not strain B-1. The band at ~30 kDa may correspond to flagellin, which has a similar predicted mass. The S-layers and flagellin are surface proteins, and may interact with the host and be recognised by the immune system.

It is difficult to distinguish between 1D gel bands, so 2D Western blotting was carried out to allow immunoreactive protein profiles to be matched to the 2D proteome reference map for each strain. Due to the large size of the 2D membranes (20x20cm), incubation with primary antibody in staining trays necessitated a much larger volume of liquid (~30ml) than the 1D blots, requiring a large amount of serum. It was also difficult to ensure incubation of the whole membrane, as the liquid tended to concentrate in the centre of the tray. Therefore, different incubation methods were tried, in order to minimise the amount of serum required for each blot, but maximise binding of the primary and secondary antibodies.

Hybridization tubes required only ~10ml antibody solution, but neither stand-alone tubes nor those rotating in a hybridisation oven gave consistent antibody binding patterns. However, hybridisation bags only required marginally more

antibody solution (~15ml), and, when incubated with rotation, consistently gave good antibody binding profiles. Therefore, hybridisation bags were used as the method of choice.

For each strain, four reproducible Western blots were selected. The SameSpots software (Non-Linear) was used to define the protein spots and align the Cy2 labelled total protein images between the four blots. Each spot was manually examined in turn to determine whether it was detected in the Cy5 (secondary antibody) channel. Immunoreactive spots were then matched to the reference map picking gel for identification.

The four Western blots used for immunoreactive protein identification in strain A are shown in *Figure 3.32* below. A total of 565 spots were identified across all the images. Twenty-one spots were manually edited to better match the protein patterns, and 24 spots were tagged as immunoreactive. These spots were then matched to the reference map picking gel using the SameSpots algorithms for identification (*Figure 3.33*).

Thirteen of the 24 immunoreactive spots were matched to identifications from the 2D reference maps, as ten different immunoreactive proteins (*Table 3.16)*

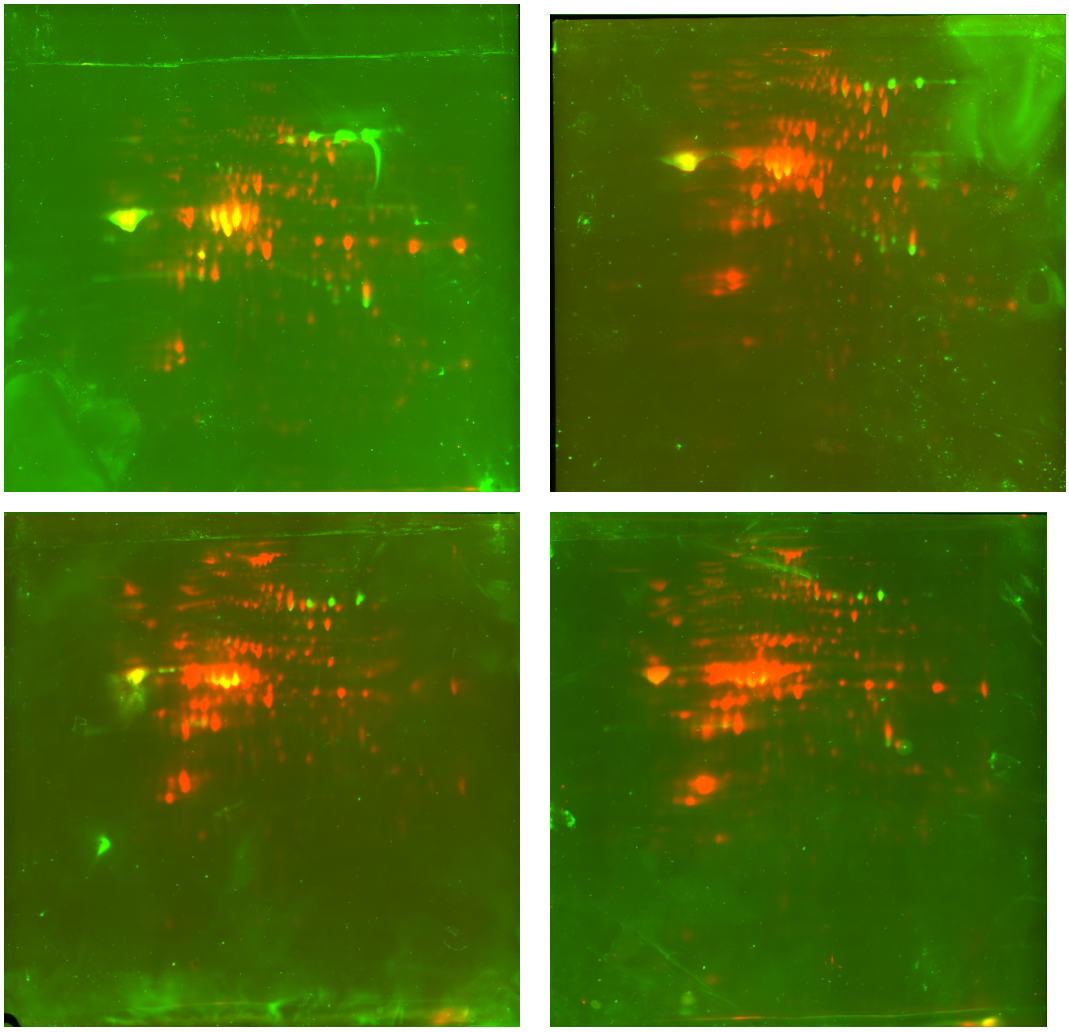


Figure 3.32(i) *– The four repeat 2D Western blots used in statistical analysis of the immunoreactive proteins detected in strain A. Total Cy2 labelled protein is shown in red, Cy5 labelled secondary antibodies in green, and overlaps appear yellow. The same pattern of clearly immunoreactive protein spots can be seen in each image, and these are shown in more detail in Figure 3.32 (ii).*

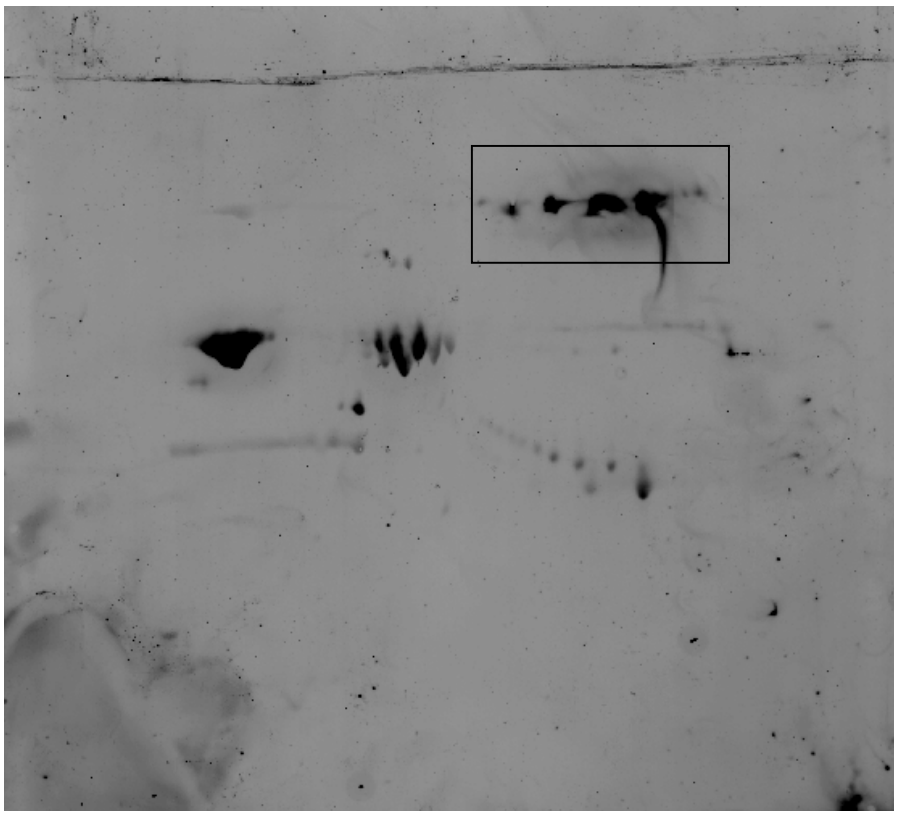


Figure 3.32(ii) *–One 2D Western blotting image showing the immunoreactive proteins detected in strain A by binding of Cy5 labelled secondary antibody. A notable set of immunoreactive protein spots detected in strain A and Tra5/5 but not B-1 is highlighted by the box.*

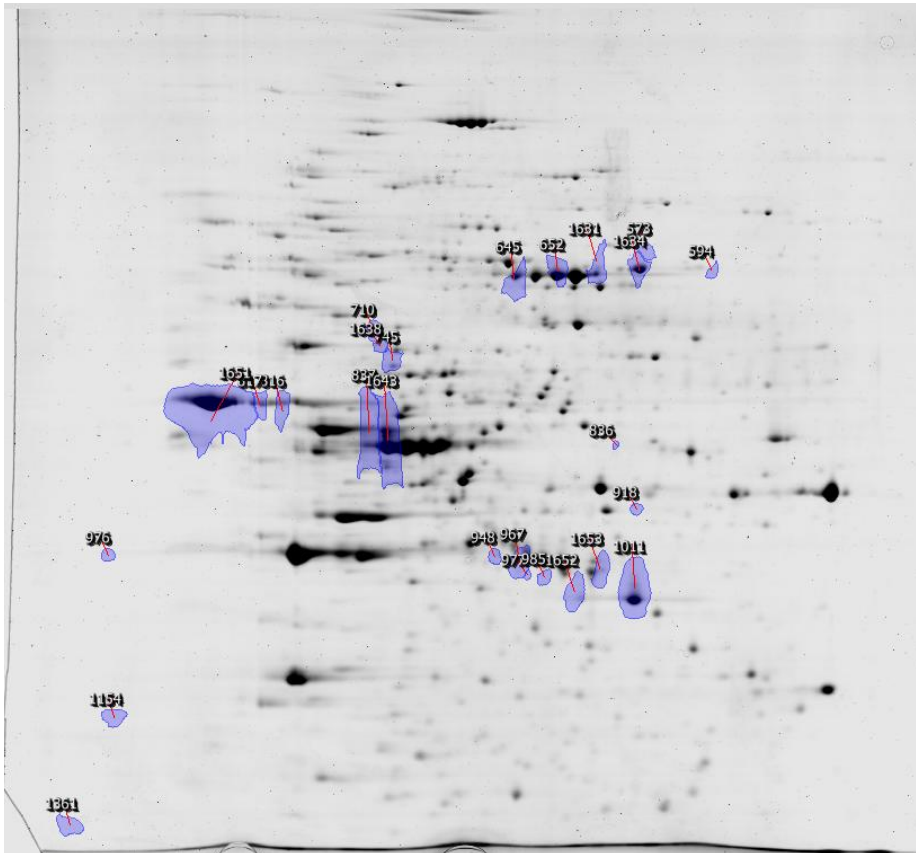


Figure 3.33 *– Proteins highlighted as immunoreactive in strain A were matched by the Samespots software to the reference map picking gel for identification. Identifications are shown in Table 3.16. Blue areas indicate the placement of the immunoreactive spots, and the identification numbers assigned by the Samespots software, based on rank, are shown.*

Table 3.16 *– Identification of the immunoreactive protein spots in strain A by alignment to the reference map.*

Anova scores and fold differences are shown for each immunoreactive protein spot, and where spots were successfully aligned, the reference map spot and protein identification is shown.

Rank	Anova (p)	Fold	Map spot	Identification number	Protein	MW	pI
1631	2.163e-006	16.5	J11	none	none		
1634	2.013e-005	12.6	J12	sta_BASYS03267	BASYS03267, 3473846-3471975 (CounterClockwise) Hypothetical Protein BASYS03267 (cell surface protein (putative S-layer protein precursor))	66427	6.2
836	1.800e-004	14.2	-				
594	2.570e-004	38.3	-				
1154	6.039e-004	19.1	-				
745	6.686e-004	3.1	-				
976	6.818e-004	17.6	-				
1361	6.836e-004	8.4	-				
573	0.002	33.5	-				
1643	0.007	2.3	E6	none	none		
			E2	sta_BASYS00306	yjiR, 302052-300856 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	44997	5.12
			E8	sta_BASYS02042	yceH, 2127930-2129066 (Clockwise) Uncharacterized protein yceH (putative tellurite resistance protein)	43643	5.09
837	0.012	2.0	E1	sta_BASYS00306	yjiR, 302052-300856 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	44997	5.12
			D2	sta_BASYS00396	(tuf1, 402179-403372 (Clockwise) Elongation factor Tu (elongation factor TU)	44226	4.94
652	0.048	4.6	J2	sta_BASYS01060	BASYS01060, 1118805-1120724 (Clockwise) Hypothetical Protein BASYS01060 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	69693	5.5
918	0.059	8.5	-				
977	0.188	3.1	-				
967	0.241	1.3	M6	sta_BASYS00598	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	34354	6.51
710	0.246	1.7	D5	none	none		
			D6	none	none		

985	0.360	2.1	M5	sta_BASYS00598	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin)	34354	6.51
1638	0.439	1.1	D7	none	none		
1651	0.456	1.5	A1	sta_BASYS03269	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 (S-layer)	80379	4.83
1652	0.495	1.3	M11	630_gi|115249239|emb|CAJ67052.1|	putative flagellar hook-associated protein	48018	4.94
1653	0.532	1.7	M2	sta_BASYS00598	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	34354	6.51
645	0.698	1.0	J4	sta_BASYS01060	BASYS01060, 1118805-1120724 (Clockwise) Hypothetical Protein BASYS01060 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	69709	5.5
817	0.778	1.7	A1	sta_BASYS03269	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 (S-layer)	80379	4.83
1011	0.886	1.1	M1	sta_BASYS01346	paaH, 1444341-1445186 (Clockwise) Hypothetical Protein paaH (3-hydroxybutyryl-CoA dehydrogenase)	30887	5.8
948	0.899	1.1	-				
816	0.931	1.5	A1	sta_BASYS03269	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 (S-layer)	80379	4.83

Four reproducible Western blots for strain B-1 were used for SameSpots analysis (*Figure 3.34*). A total of 461 spots were identified across the images which were manually examined to see if they showed any immunoreactivity. Eleven spots were edited, and 43 were tagged as immunoreactive. These spots were then matched to the reference map picking gel (*Figure 3.35*). Twenty-eight of the 43 immunoreactive spots were identified from the 2D reference maps as 18 different immunoreactive proteins (*Table 3.17*).

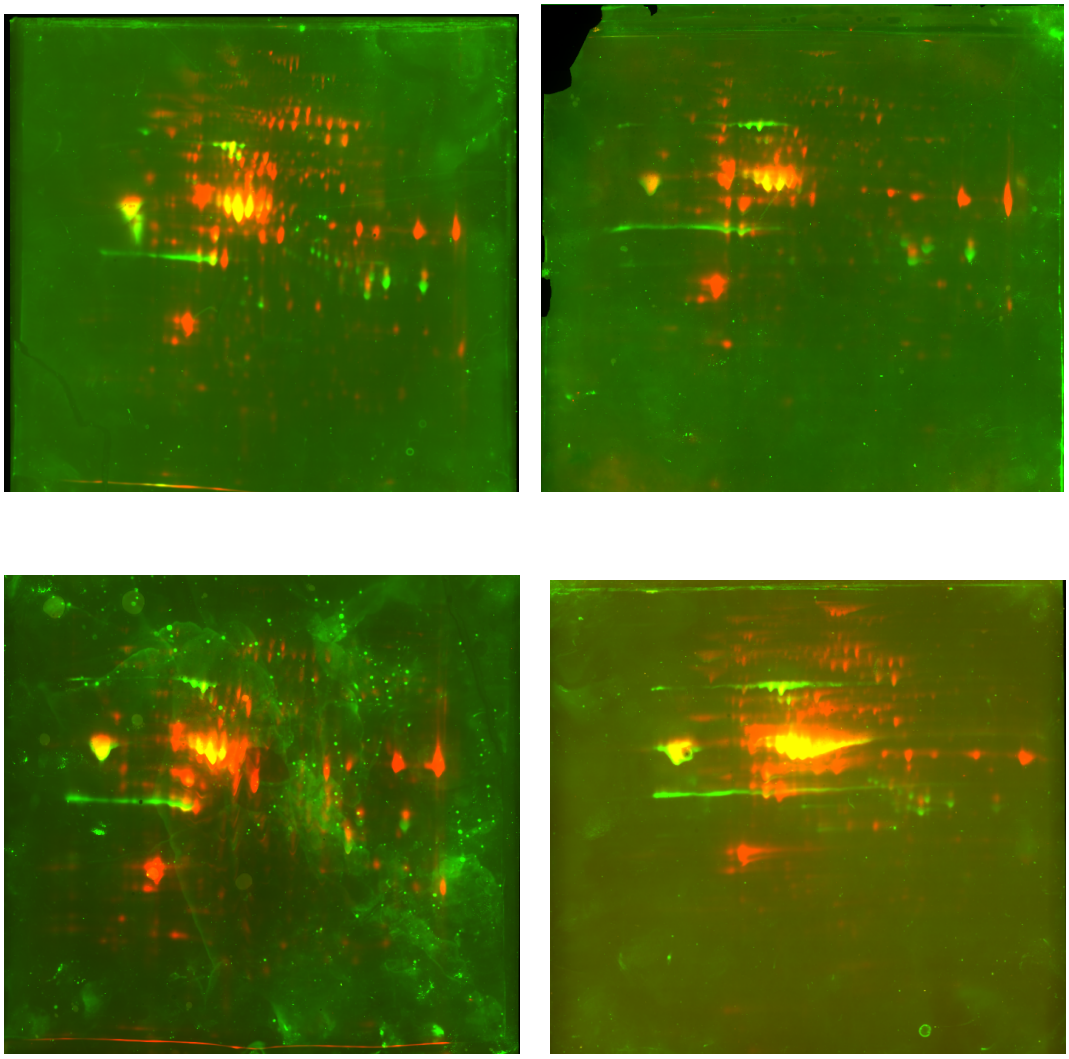


Figure 3.34(i) *– The four repeat 2D Western blots used in statistical analysis of the Immunoreactive proteins detected in strain B-1. Total Cy2 labelled protein is shown in red, Cy5 labelled secondary antibodies in green, and overlaps appear yellow. Again the pattern of immunoreactive spots is reproducible and is shown in more detail in (ii).*

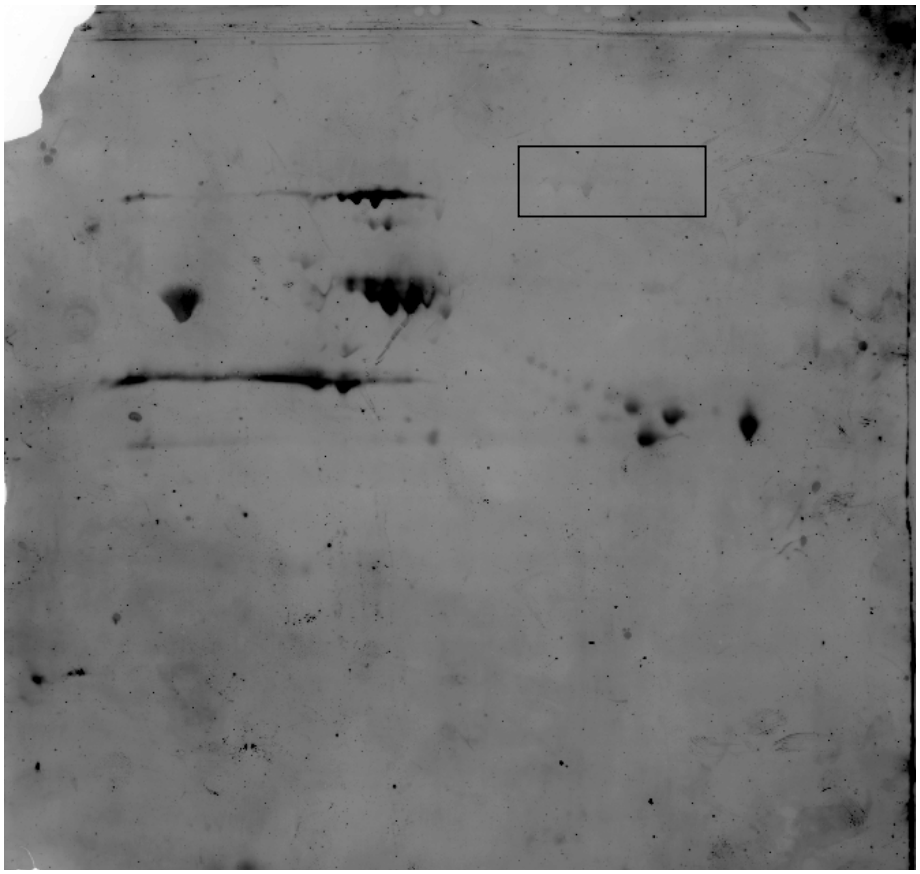


Figure 3.34(ii) *– One 2D Western blotting image showing the immunoreactive proteins detected in strain B-1 by binding of Cy5 labelled secondary antibody. The area lacking the notable set of spots detected in strain A and Tra5/5 but not B-1 is highlighted by the box.*

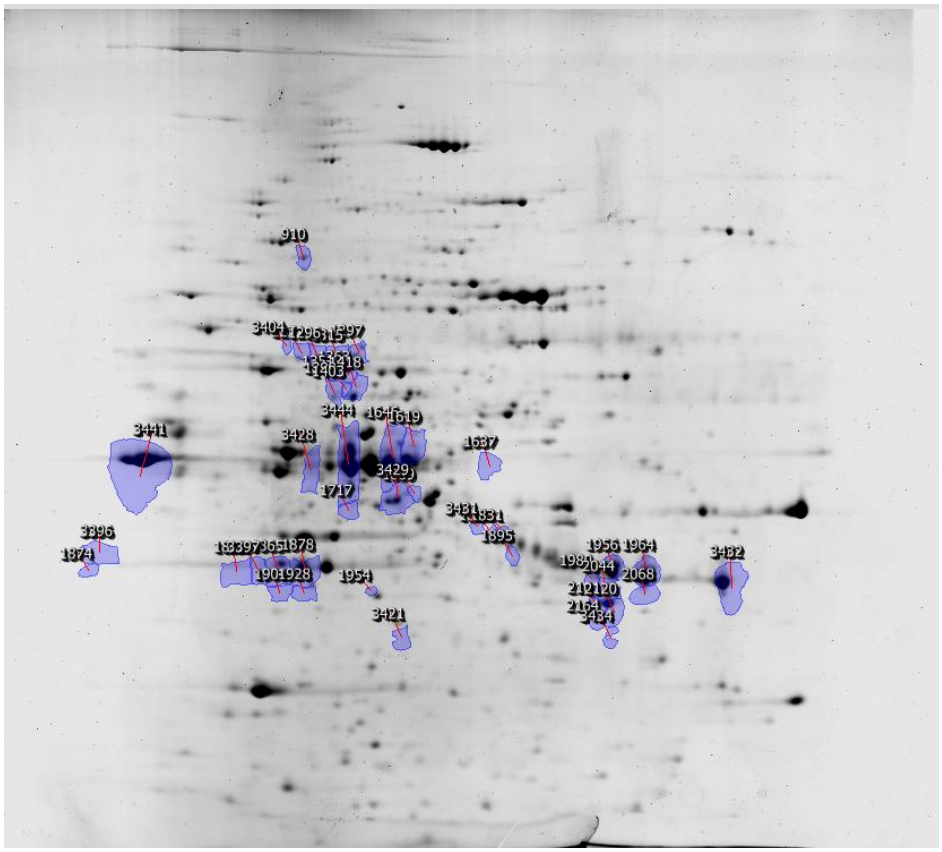


Figure 3.35 *– Proteins highlighted as immunoreactive in strain B-1 were matched by the SameSpots software to the reference map picking gel for identification. Identifications are shown in Table 3.17. Blue areas indicate the placement of the immunoreactive spots, and the identification numbers assigned by the Samespots software, based on rank, are shown.*

Table 3.17 *– Identification of the immunoreactivve protein spots in strain B-1 by alignment to the reference map.*
Anova scores and fold differences are shown for each immunoreactive protein spot, and where spots were successfully aligned, the reference map spot and protein identification is shown.

Rank	Anova (p)	Fold	Map spot	Identification number	protein	MW	pI
1271	1.124e-007	7.3	D5	stb_BASYS00533	oligopeptide ABC transporter, substrate-binding lipoprotein	58649	5.43
1862	1.306e-006	8.1	-				
1296	4.242e-006	7.5	D6	stb_BASYS00533	oligopeptide ABC transporter, substrate-binding lipoprotein	58649	5.43
3397	2.077e-005	6.6	F4	stb_BASYS02565	thioredoxin reductase	34471	4.91
1315	8.533e-005	6.8	D6	stb_BASYS00533	oligopeptide ABC transporter, substrate-binding lipoprotein	58649	5.43
			D7	stb_BASYS02898	oligopeptide ABC transporter, substrate-binding protein	58299	5.36
3396	1.531e-004	19.7	-				
3404	5.696e-004	4.6	-				
3421	0.001	2.8	-				
1874	0.002	27.5	-				
1297	0.004	7.3	-				
1865	0.010	2.7	D4	stb_BASYS01043	rpsA, 1080759-1082033 (Clockwise) Hypothetical Protein rpsA (putative 30S ribosomal protein S1)	48454	5.02
1418	0.022	1.6	D10	stb_BASYS01667	aegA, 1712608-1714002 (Clockwise) Hypothetical Protein aegA (putative glutamate synthase NADPH small chain)	50234	5.12
3429	0.034	1.8	E9	stb_BASYS02338	elongation factor Ts	33290	5.21
			E10	none	none		
1646	0.040	2.0	E4	stb_BASYS00983	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	44491	5.18
1831	0.042	5.7	M9	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1954	0.044	4.6	F6	stb_BASYS00349	flagellar hook protein	34568	5.24
1980	0.045	3.9	-				
2044	0.049	1.8	M12	stb_BASYS01116	3-hydroxybutyryl-CoA dehydrogenase	30887	5.8
			M4	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1895	0.061	2.2	-				

Rank	Anova (p)	Fold	Map spot	Identification number	protein	MW	pI
2164	0.070	5.5	M12	stb_BASYS01116	3-hydroxybutyryl-CoA dehydrogenase	30887	5.8
1807	0.087	5.1	M10	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1878	0.091	1.8	F5	stb_BASYS00514	ABC transporter, substrate-binding lipoprotein	36159	5.4
1637	0.099	2.1	-				
1928	0.099	1.1	-				
1403	0.115	2.5	-				
3434	0.117	3.4	-				
1619	0.141	2.0	E5	stb_BASYS01117	acetyl-CoA acetyltransferase	41005	5.2
			F1	stb_BASYS00980	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	42966	5.27
2068	0.163	2.6	-				
3432	0.172	2.2	M1	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1964	0.173	2.7	M3	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1956	0.187	2.6	M4	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1351	0.200	2.9	D6	stb_BASYS00533	oligopeptide ABC transporter, substrate-binding lipoprotein	58649	5.43
			D7	stb_BASYS02898	oligopeptide ABC transporter, substrate-binding protein	58299	5.36
910	0.217	1.9	C10	stb_BASYS02649	glyS, 3040737-3038671 (CounterClockwise) Hypothetical Protein glyS (glycyl-tRNA synthetase beta chain)	78491	4.97
3431	0.223	4.7	M10	stb_BASYS00368	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	34339	7.82
1363	0.234	3.3	D7	stb_BASYS02898	oligopeptide ABC transporter, substrate-binding protein	58299	5.36
2101	0.234	1.8	M12	stb_BASYS01116	3-hydroxybutyryl-CoA dehydrogenase	30887	5.8
1717	0.277	1.1	E8	stb_BASYS01826	putative tellurite resistance protein	43646	5.09
3441	0.326	1.4	A1	stb_BASYS03072	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	64664	4.76
1904	0.338	1.3	-				
1399	0.426	1.1	-				
3444	0.456	1.4	E6	stb_BASYS00210	NAD-specific glutamate dehydrogenase	46271	5.16
			E6	stb_BASYS04027	putative amino acid aminotransferase	44997	5.12
			E6	stb_BASYS01452	aspartate aminotransferase	45066	5.1
			E2	stb_BASYS04027	putative amino acid aminotransferase	44997	5.12
2120	0.501	1.0	M12	stb_BASYS01116	3-hydroxybutyryl-CoA dehydrogenase	30887	5.8
1689	0.628	1.6	-				
3428	0.820	1.1	E1	stb_BASYS04027	putative amino acid aminotransferase	44997	5.12

The four strain Tra5/5 Western blots used in SameSpots analysis are shown in *Figure 3.36* below. A total of 421 spots were matched across the images and examined for any immunoreactivity. 58 spots were manually edited to better fit the 2D protein profile, and 33 spots were tagged as immunoreactive and matched to the reference map picking gel (*Figure 3.37*). Twenty-three of 33 immunoreactive spots were identified from the 2D reference maps as a total of 18 different proteins.

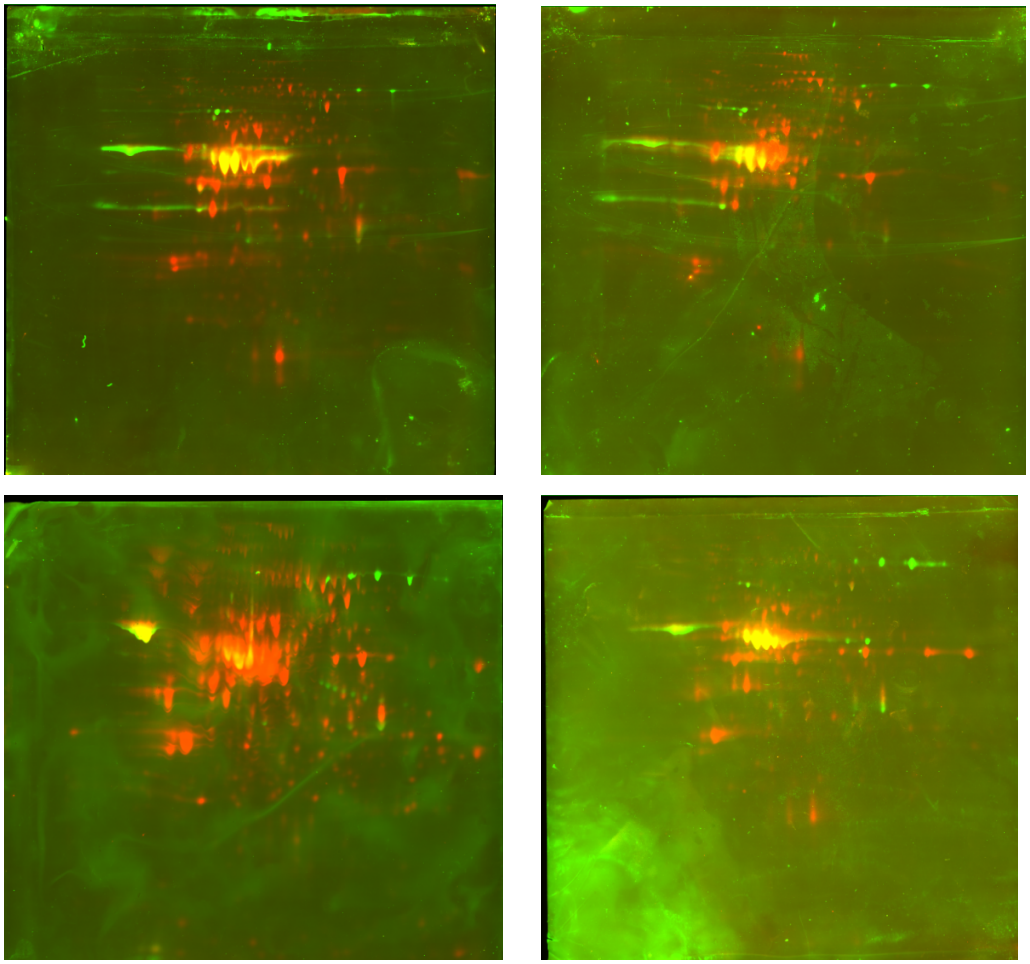
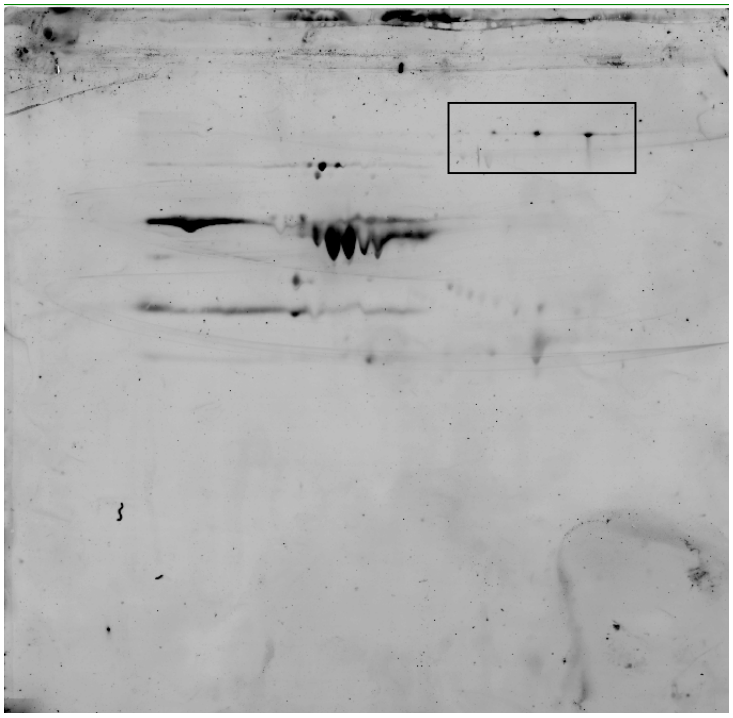


Figure 3.36(i) *– The four repeat 2D Western blots used in statistical analysis of the Immunoreactive proteins detected in strain Tra5/5. Total Cy2 labelled protein is shown in red, Cy5 labelled secondary antibodies in green, and overlaps appear yellow. Immunoreactive proteins are shown in detail in (ii).*



.

Figure 3.36(ii) *– One 2D Western blotting image showing the immunoreactive proteins detected in strain Tra5/5 by binding of Cy5 labelled secondary antibody. Again, the series of immunoreactive spots present in strains A and Tra5/5 but not B-1 is highlighted within the box.*

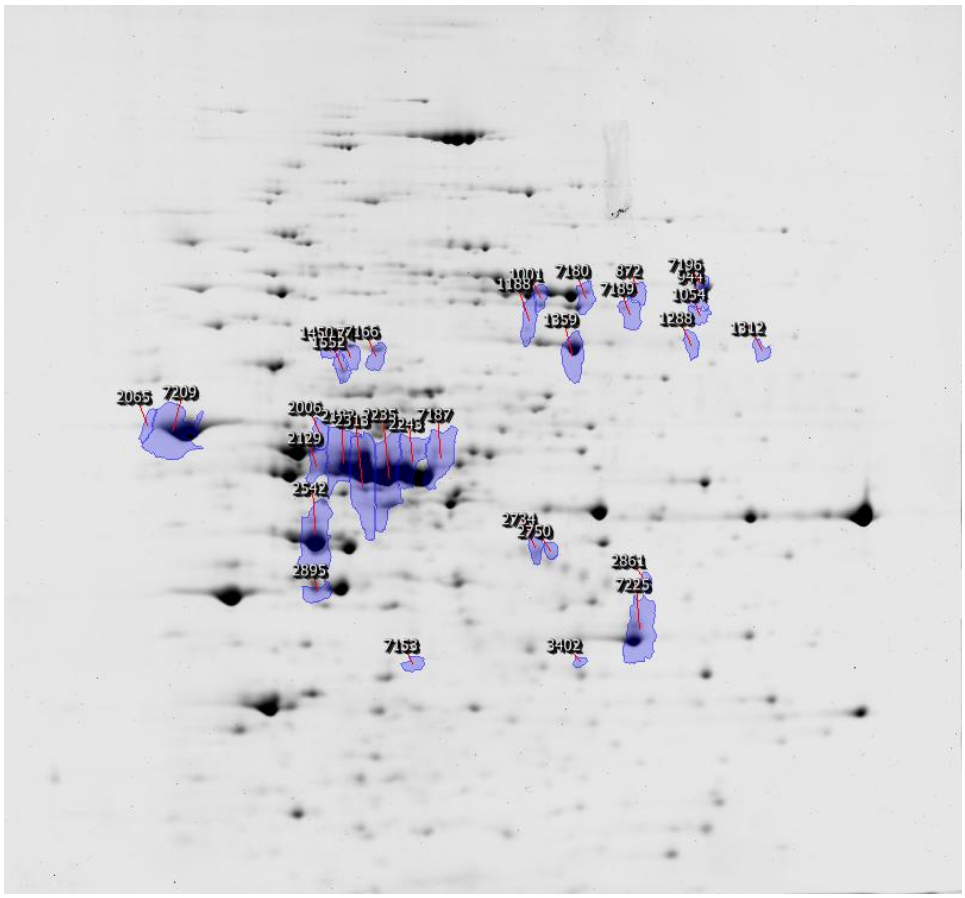


Figure 3.37 *– Proteins highlighted as immunoreactive in strain Tra5/5 were matched by the SameSpots software to the reference map picking gel for identification. Identifications are shown in Table 3.18. Blue areas indicate the placement of the immunoreactive spots, and the identification numbers assigned by the Samespots software, based on rank, are shown.*

Table 3.18 *– Identification of strain Tra5/5 immunoreactive proteins from reference maps. Anova scores and fold differences are shown for each immunoreactive protein spot, and where spots were successfully aligned, the reference map spot and protein identification is shown.*

Rank	Anova (p)	Fold	Map spot	Identification number	protein	MW	pI
2542	2.215e-006	6.3	A5	stt_BASYS00466	fixB, 450899-451936 (Clockwise) Hypothetical Protein fixB (electron transfer flavoprotein alpha-subunit)	57353	5.03
1359	1.328e-004	4.2	K1	stt_BASYS00861	BASYS00861, 866100-867794 (Clockwise) Hypothetical Protein BASYS00861 (formate--tetrahydrofolate ligase)	60894	5.53
7153	4.216e-004	3.6	-				
2243	6.695e-004	4.6	E4	stt_BASYS00460	frc, 444475-445674 (Clockwise) Hypothetical Protein frc (isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase)	44491	5.18
7180	0.001	5.1	J11	stt_BASYS02456	BASYS02456, 2546072-2544243 (CounterClockwise) Hypothetical Protein BASYS02456 (indolepyruvate oxidoreductase subunit)	68911	5.65
7187	0.002	2.8	F1	stt_BASYS01130	atoB, 1171926-1173158 (Clockwise) Hypothetical Protein atoB (acetyl-CoA acetyltransferase)	43350	5.56
			E5	stt_BASYS01130	atoB, 1171926-1173158 (Clockwise) Hypothetical Protein atoB (acetyl-CoA acetyltransferase)	43350	5.56
1188	0.002	4.7	J3	stt_BASYS00858	BASYS00858, 862861-864780 (Clockwise) Hypothetical Protein BASYS00858 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	69693	5.5
2734	0.007	4.9	M8	stt_BASYS00250	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	34339	7.82
7192	0.012	3.3	-				
2006	0.012	1.7	-				
1312	0.021	4.6	-				
1054	0.021	7.4	J12	stt_BASYS02896	BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (cell surface protein	66363	6.47
1288	0.024	13.6	-				
2182	0.024	1.3	E1	stt_BASYS03814	yjiR, 4076689-4075493 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	44997	5.12
7189	0.038	7.4	-				

Rank	Anova (p)	Fold	Map spot	Identification number	protein	MW	pI
2065	0.045	2.2	A1	stt_BASYS02898	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898 (S-layer)	80379	4.83
872	0.051	21.6	J10	none			
7196	0.055	24.7	J12	stt_BASYS02896	BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (cell surface protein (putative S-layer protein precursor))	66363	6.47
1474	0.056	4.9	P11	none			
944	0.058	31.0	J12	stt_BASYS02896	BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (cell surface protein (putative S-layer protein precursor))	66363	6.47
2313	0.065	1.5	E6,E2,E8	stt_BASYS03814	yjiR, 4076689-4075493 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	44997	5.12
			E2	stt_BASYS03814	yjiR, 4076689-4075493 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	44997	5.12
			E8	stt_BASYS02589	hisC, 2688605-2689708 (Clockwise) Hypothetical Protein hisC (putative histidinol-phosphate aminotransferase)	42016	5.09
1450	0.065	3.1	-				
7209	0.072	1.9	A1	stt_BASYS02898	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898	80379	4.83
2750	0.089	2.9	M7	stt_BASYS00250	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	34339	7.82
7225	0.121	1.5	N5	stt_BASYS01129	paaH, 1171051-1171896 (Clockwise) Hypothetical Protein paaH (3-hydroxybutyryl-CoA dehydrogenase)	30887	5.8
			N6	stt_BASYS01129	paaH, 1171051-1171896 (Clockwise) Hypothetical Protein paaH (3-hydroxybutyryl-CoA dehydrogenase)	30887	5.8
2235	0.134	1.4	E3	stt_BASYS00188	gluD, 189677-190942 (Clockwise) NAD-specific glutamate dehydrogenase	46271	5.16
1001	0.225	3.6	J2				
			J3	stt_BASYS00858	BASYS00858, 862861-864780 (Clockwise) Hypothetical Protein BASYS00858 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	69693	5.5
2129	0.434	1.1	E1	stt_BASYS03814	yjiR, 4076689-4075493 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	44997	5.12
3402	0.469	2.1	-				

7166	0.496	1.1	P10	stt_BASYS03698	lysS, 3955558-3954029 (CounterClockwise) Lysyl-tRNA synthetase (lysyl-tRNA synthetase)	58539	5.12
2895	0.652	1.5	F5	stt_BASYS00458	ldhA, 443798-442800 (CounterClockwise) Hypothetical Protein ldhA ((R)-2-hydroxyisocaproate dehydrogenase)	36552	5.07
2861	0.850	1.3	M3	stt_BASYS00250	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	34339	7.82
1552	0.956	1.6	D5	stt_BASYS01605	aegA, 1647197-1648591 (Clockwise) Hypothetical Protein aegA (putative glutamate synthase NADPH small chain)	50248	5.12
			D6	stt_BASYS02768	yddS, 2891711-2893306 (Clockwise) Hypothetical Protein yddS (oligopeptide ABC transporter, substrate-binding protein)	58299	5.51

3.4.3 – Human Serum

Once a reliable platform for Western blotting with rabbit serum had been established, Western blotting was carried out with human patient sera to try to determine differences in the immune responses of patients with infections of varying severity against the three strains.

Pooled serum samples from patients were obtained from Professor Ian Poxton at the University of Edinburgh. For each pool, five different serum samples from the following groups of patients were used.

D) Cases – symptomatic, toxin-positive and culture-positive
E) Carriers – asymptomatic, culture positive, toxin-variable
F) Controls – asymptomatic, toxin-negative, culture-negative

Controls were matched for age, and were in the same wards as patients in groups A and B.

The same protocol was used for Western blotting of the human serum samples as used for the rabbit serum samples. The secondary antibody used was a Cy5 labelled goat polyclonal to human IgG.

In order to ensure that the anti-human secondary antibody bound to the human serum at the same concentrations as the anti-rabbit secondary antibody bound to the rabbit serum, a series of dot blots were carried out at a number of dilutions of serum (*Figure 3.38*) Membranes were blocked with 1% BSA for 1 hour before incubation with either 1:10,000 or 1:20,000 dilution of secondary antibody.

1/100	1/100	1/50	1/50
1/500	1/500	1/200	1/200
1/2000	1/2000	1/1000	1/1000
1/10000	1/10000	1/5000	1/5000
1/50000	1/50000	1/20000	1/20000

Figure 3.38(i) *– Serial dilutions of serum (primary antibody) used in dot blots.*

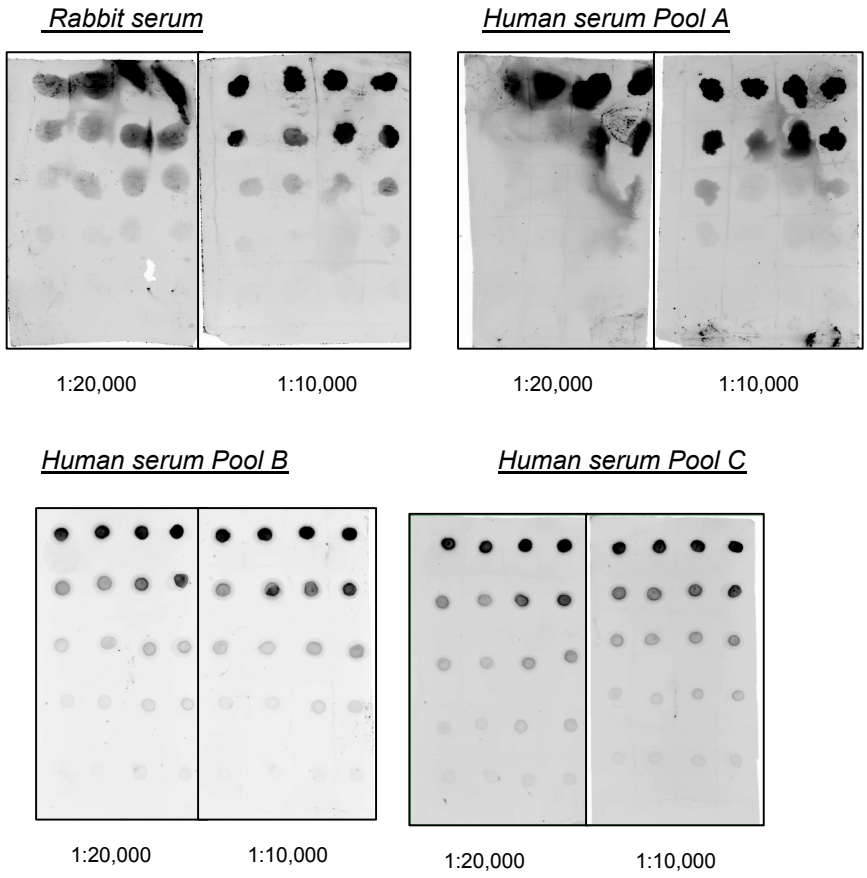


Figure 3.38(ii) *– Dot blots of each serum sample, serially diluted as in (i) and incubated with the relevant secondary antibody (Cy5 anti-rabbit or Cy5 anti-human) at two different concentrations.*

All three human serum samples showed comparable immunoreactivity to the anti-human secondary anti-body to that of the rabbit serum to anti-rabbit secondary antibody. However, probing the blots using the human serum at a 1:10,000 dilution, as used successfully for the rabbit serum, gave no results, so greater concentrations of human serum were used.

Five μg Cy2 labelled protein samples from each of the three *Clostridium difficile* strains were separated by 1D gel electrophoresis, and the proteins transferred to nitrocellulose membrane as previously.
Membranes were blocked overnight with 1% BSA, incubated with human serum at 1:1000 dilution for 2 hours, and then incubated with secondary antibody at a dilution of 1:10,000 for 2 hours.

Firstly, the pool A serum from symptomatic patients was used as the primary antibody (*Figure 3.39*).

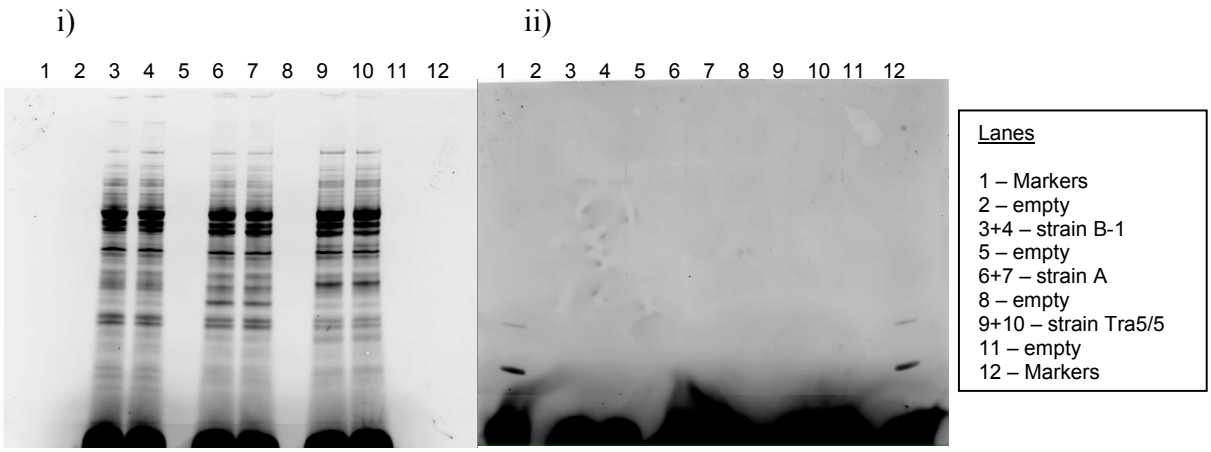


Figure 3.39 *– Western blotting of C. difficile protein extracts probed with human serum pool A. i) Total protein extract labelled with Cy2 before electrophoresis; ii) Immunoreactive proteins bound by Cy5 labelled secondary antibody.*

No *C. difficile* proteins were bound by the serum antibodies, but non-specific binding to the marker proteins (aprotinin from horse muscle and lysozyme from chicken egg) was observed. The marker proteins of smaller sizes are present at high quantity which may explain this non-specific interaction. There was also substantial straining at the gel front (*Figure 3.39*).
Incubation time, concentration of the *Clostridium difficile* protein extracts, and dilution factor of the human serum and secondary antibody were all varied, but only the same, non-specific interactions were seen.

Western blots were then carried out with the serum pool B (carriers).
Membranes were blocked overnight with 1% BSA, incubated with human serum pool B at 1:2000 dilution for 2 hours, and then incubated with secondary antibody at a dilution of 1:10,000 for 2 hours (*Figure 3.40*).

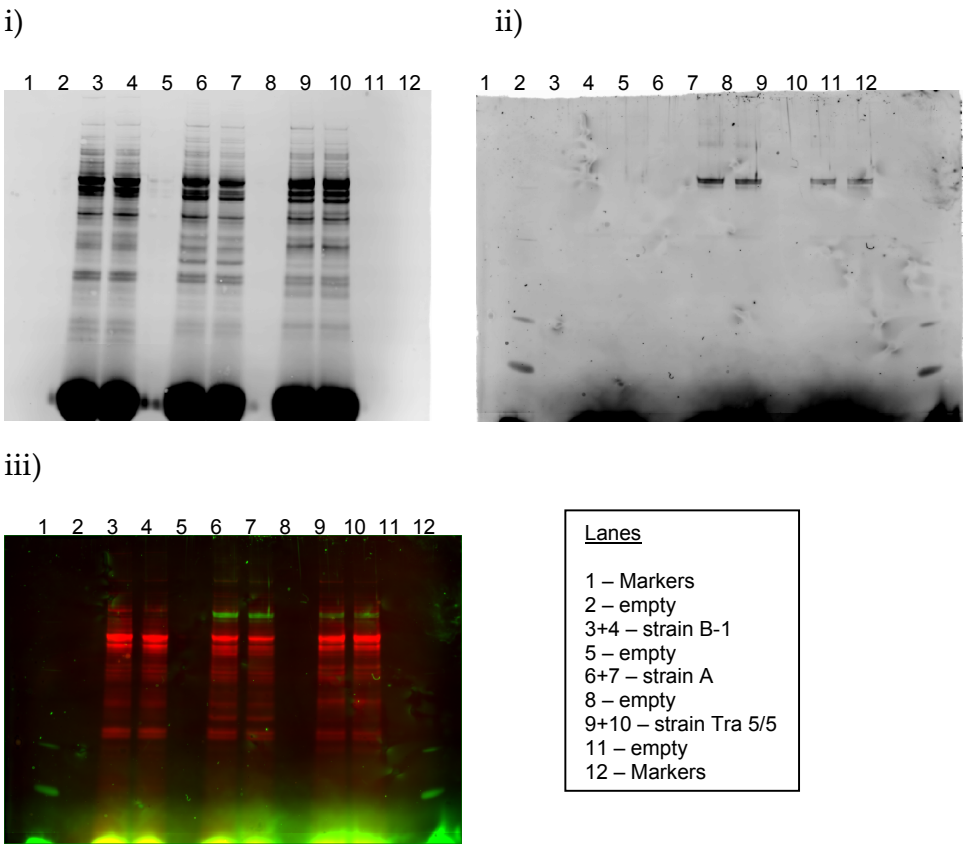


Figure 3.40 *- Western blotting of C. difficile protein extracts probed with human serum pool B. (i) Total protein extract labelled with Cy2 before electrophoresis; (ii) Immunoreactive proteins bound by Cy5 labelled secondary antibody; (iii) Both channels visualised at once.*

There was a distinct immunoreactive protein band present in strains A and Tra5/5, but not B-1 (*Figure 3.40*) which may correspond to the band of ~66 kDa seen in strains A and Tra5/5 but not B-1 in the rabbit serum blots.
However, this pattern of bands could not be repeated. Different concentrations of primary and secondary antibody were tried, along with the use of EDTA in the antibody solution, but no immunoreactive bands were detected in further blotting experiments. Again, substantial staining at the gel front was observed.

Finally, the Western blots were run with the control serum, pool C. Membranes were blocked overnight with 1% BSA, incubated with human serum pool C at 1:2000 dilution for 2 hours, and then incubated with secondary antibody at a dilution of 1:10,000 for 2 hours.

No immunoreactive bands were detected at all using the pool C serum (*Figure 3.41*), although the same staining at the gel front was observed.

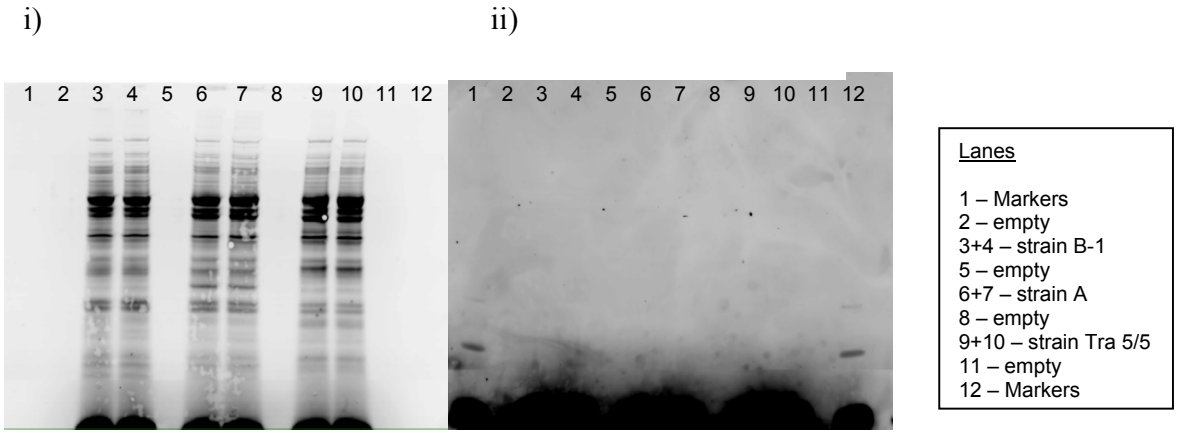


Figure 3.41 *– Western blotting of C. difficile protein extracts probed with human serum pool C. (i) Total protein extract labelled with Cy2 before electrophoresis; (ii) Immunoreactive proteins bound by Cy5 labelled secondary antibody*

Chapter 4 – Discussion

To date, the proteome of *Clostridium difficile* has not been widely characterised. Previous work has focused on single protein/enzyme analysis, most frequently on the toxins, but two studies have adopted a wider proteomic approach to characterise this organism. The cell surface proteins of strain 630 have been characterised (Wright *et al*, 2005), as has the composition of spores from that strain (Lawley *et al*, 2009). In addition, proteins released during high toxin production have been examined and compared for different strains. (Mukherjee *et al*, 2002).

These reports have highlighted a number of secreted proteins, and cell surface located proteins that are likely to be important for adhesion and bacterial – host interaction, as well as proteins involved in sporulation and germination pathways. However, no analysis of cellular protein composition, or comparative analysis of proteomic differences between isolates of varying virulence has so far been described. This study utilised a combination of gel based proteomic methodologies coupled with MALDI-TOF and LC-MS/MS mass spectrometry to compare the whole-cell proteomes of three *C. difficile* isolates of different ribotypes and virulence, and from different epidemiological backgrounds. Such gel-based proteomic techniques provide powerful tools for the separation and identification of individual components of complex protein mixtures, such as bacterial whole cell extracts, and have been reviewed as a successful strategy for use on protein samples from Gram-positive bacteria by Hecker *et al* (Antelmann and Hecker, 2010; Hecker *et al*, 2008).

4.1 – Growth Curves and optimisation of Protein Extraction

Unlike the static genome, the proteome of an organism is dynamic and will vary enormously depending on growth phase and environmental conditions. Therefore, when undertaking comparative proteomic analysis of different strains, it is essential to control for these variables as much as possible, to allow true strain-to-strain differences to be detected. In addition, it is desirable to maximise the coverage of the proteome so a greater proportion of proteins can be compared. Hence the preparation of the whole cell protein extracts is of the utmost importance, and optimisation is key.

Plate cultures grown for 64 hours and extraction using the Fast-Prep system gave extensive and reproducible profiles as well as being a simple and straightforward protocol, so these were used as the standard extracts (*Figure 3.5*). When extracting from plate cultures it cannot be determined exactly what phase of growth cells are in, and a wide range of cells in different growth phases are likely to be present. This may explain why plate cultures gave more extensive profiles.

For this project, the standard protein extracts aimed to maximise the proportion of the proteome detected, hence the use of plate cultures. The three strains were found to grow and progress through the growth cycle at comparable rates, so cultures incubated for the same length of time under the same conditions could be considered to be in the same phase.

It would be expected that after 64 hours of growth, some sporulation would occur. However, the protein extraction techniques were optimised for vegetative cells, and the standard plate extracts show similarity to published whole cell protein profiles of vegetative cells, and little similarity to protein extracts from purified spores (Lawley *et al*, 2009), indicating the vegetative proteome is dominant. Sporulation proteins identified were nevertheless examined in detail, to see if differences in the sporulation timing or mechanism were indicated between strains.

4.2 – Protein Identification

4.2.1 – Reference mapping

Although 2D reference mapping identifies only 10 – 30% of the most abundant proteins in the proteome (Cordwell *et al*, 2001), it has been the methodology of choice for proteomic analysis of a large range of bacteria (Encheva *et al*, 2005); (Bernardini *et al*, 2004) (Riedel *et al*, 2006); (Eymann *et al*, 2004), including the clostridial species *C. perfringens* (Alam *et al*, 2009) and *C. acetobuylicum* (Sullivan and Bennett, 2006). Reference mapping is especially useful in identifying differences between protein profiles.

Between 80 and 120 proteins were identified per strain using MALDI-TOF MS, but only a low proportion (29%) of all identified proteins were detected in all three strains, the majority of which were abundant metabolic enzymes. Many proteins were identified in one strain only, including a number of potential virulence factors such as sporulation, flagellar and S-layer proteins, with a

greater proportion of these proteins of interest detected in the highly virulent strain A than the other two strains.

Reference mapping using MALDI-TOF provides a visual profile of many components of the proteome, but has some drawbacks. Protein separation is key, as MALDI does not easily identify mixtures of proteins. A maximum of about three proteins can be identified from one sample. MALDI is also only useful in identifying highly abundant proteins, hence the small proportion of the proteome detected here.

It is difficult to determine whether observed strain to strain difference reflects true expression difference, or is partially a consequence of the difficulties associated with reference mapping, so additional analysis using LC-MS/MS was used to increase coverage of the proteome and confirm expression differences.

4.2.2 – LC-MS/MS

LC-MS/MS is a more sensitive technique which allows identification of many proteins from a complex mixture, so can be used with 1D GE instead of requiring 2D electrophoresis, and can also detect proteins present at much lower concentrations.

LC-MS/MS analysis greatly improved the coverage of the proteome, with more than 80% of the proteins identified by this method not detected by MALDI-TOF MS. In particular, many additional proteins with theoretical mass and pI values outside the range of the 2D gel electrophoresis were identified. However,

almost 40% of the proteins identified were still detected in one strain only, indicating that there are expression differences between the strains.

Both LC-MS/MS and reference mapping indicated that a greater proportion of potential virulence factors were present in strain A than in the other two strains, suggesting that the reported hypervirulence of this strain may have a proteomic basis. A number of additional sporulation and cell surface proteins were detected in strain A but not the other two strains, as well a several putative membrane proteins. A number of flagellar proteins were detected only in strain B-1, and only one potential virulence factor, a cell surface protein, was identified only in strain Tra5/5.

Given that strains A and B-1 have both been shown to be highly virulent, but strain Tra5/5 is a low virulence strain, proteins identified in strains A and B-1 but not Tra5/5 are of interest. Most notably, these include toxin B, one of the main *Clostridium difficile* disease causing agents. The flagellar hook protein and flagellar basal body protein were also not detected in Tra5/5 indicating differences in flagellar structure. A number of ABC transporter proteins were found in A and B-1 but not Tra5/5, as were four cell surface proteins, all of which showed high homology between strains.

In total, 16.50% of the predicted ORFs from strain A, 16.65% of the predicted ORFs in strain B-1, and 11.64% of the predicted ORFs in strain Tra5/5 have been confirmed in this study by protein identification. This is a larger proportion of identifications than has often been reported based on 2D reference mapping alone (Encheva *et al*, 2005); (Riedel *et al*, 2006), and a similar proportion of

identifications to that reported for similar, whole cell protein extracts in *Bacillus subtilis* (21%) (Eymann *et al*, 2004) and *Neisseria meningitidis* (13%) (Bernardini *et al*, 2007b).

Despite the increased sensitivity of LC-MS/MS identification, it should be noted that there was still a subset of 26 proteins identified by reference mapping that were not detected by LC-MS/MS, highlighting the importance of using a range of techniques to increase coverage of the bacterial proteome.
However, the use of 1D GE in conjunction with LC-MS/MS limits identification to those proteins compatible with gel electrophoresis. To improve proteome coverage further, gel free identification methods should be considered.

4.2.3 – DIGE analysis

Although many proteins have been identified on the reference maps, direct comparison of 2D gel profiles is problematic. Differences in the amount of protein loaded, and individual variation in the running of the gels can complicate gel to gel comparisons, making meaningful evaluation of these reference maps difficult.
Differential in-gel electrophoresis allows different protein samples to be run on a single gel, eliminating the experimental differences and allowing direct comparison (Unlu *et al*, 1997), and was used here to analyse the 2D reference maps. DIGE can also be coupled with 2D reference mapping to glean quantitative information for proteins with respect to an internal reference sample (Hecker *et al*, 2008); (Encheva *et al*, 2009); (Rathsam *et al*, 2005), as done in this analysis where a number of proteins showing statistically significant differences in expression levels between strains were identified.

Five proteins were shown to have statistically higher concentration in strain A, twelve in strain B-1 and eight in strain Tra5/5. A number of these were surface proteins, with different S-layer proteins found have higher concentration in each strain, and the flagellin protein showing statistically greater abundance in both A and B-1.

The S-layer protein with the highest homology to the *SlpA* gene product from the 630 reference strain (gi|115251846), showed significantly different expression between the three isolates investigated here. However, as highlighted in the yellow box 4, *Figure 3.13*, this protein migrated differently in different samples. It has previously been reported that the S-layer proteins of different isolates show a high degree of variation (McCoubrey and Poxton, 2001); (Calabi *et al*, 2001); (Pechine *et al*, 2005b) and the genomic data for the three strains in this study indicates that this is true of the *SlpA* genes of A, B-1 and Tra5/5. The large divergence in amino acid sequence of this protein leads to different predicted mass and pI values for each strain, with predicted pI values of 4.83, 4.76 and 4.79, and predicted masses of 80428 Da, 64703 Da and 79922 Da for the S-layer protein of A, B-1 and Tra5/5 respectively. Therefore, the concentration differences highlighted by the software are likely to be due to differences in migration rather than true variation in expression.

Interestingly, this protein is present as two spots on the reference maps of strains A (A1, F6, *Figure 8*) and Tra5/5 (A1, F6, *Figure 10*), but only one on the map of strain B-1 (A1, *Figure 9*), indicating that there are modification differences between the strains. The *SlpA* gene product is cleaved post-translationally to give two mature proteins of ~45kDa (the high molecular

weight S-layer protein) and ~35kDa (the low molecular weight S-layer protein) (Calabi *et al*, 2001). The positions of the SlpA homologous proteins on the 2D profiles correspond to these approximate sizes, and similar 2D profiles of S-layer proteins have been shown previously (Mukherjee *et al*, 2002). Here, the spot corresponding to the low molecular weight S-layer protein (LMW SLP) is missing from the profile of strain B-1. The LMW SLP has also been shown by DIGE to be upregulated in both strains A and Tra5/5 with respect to each other, as well as with respect to B-1, but again this is likely to be due to amino acid variation causing migration differences.

4.3 – ProteoMiner treatment

The identification strategies discussed so far have a major drawback, in that they can only detect a fairly small proportion of the proteome. The huge dynamic range of proteins in a highly complex sample such as the *C. difficile* cell extracts used here, mean that proteins present at low abundance may not be detected. In order to try to increase coverage of this 'deep proteome', samples were treated with a hexapeptide bead library, the ProteoMiner beads (BioRad). ProteoMiner has been successfully used on a number of biological fluids (Guerrier *et al*, 2007); (Castagna *et al*, 2005); (Sihlbom *et al*, 2008), but has not been widely used on bacterial extracts, although some success has been reported with *E. coli* whole cell extracts (Thulasiraman *et al*, 2005).

One of the challenges of using this technology on bacterial extracts is the large amount of protein required to sufficiently decrease the dynamic range of the sample. It was impractical to generate such large amounts of protein extracts

from plate cultures as was done for the standard protein extracts, so large scale broth cultures were used instead.

Environmental conditions, and hence growth media, have a significant effect on protein expression (Cordwell *et al*, 2001). Here, more than 50% of the proteins identified by LC-MS/MS in the standard plate extracts from each strain were not detected in the broth extracts used for ProteoMiner treatment (57% in strain A, 64% in strain B-1 and 55% in strain Tra5/5). In addition, approximately 50% of the proteins identified in the broth extracts were not detected in the standard plate extracts (54% for strain A, 48% for strain B-1 and 66% for strain Tra5/5). This underlines the fact that consistency of growth conditions is key when comparing protein samples, and that the additional investigation of extracts from cultures grown in different media can greatly increase the coverage of the proteome.

However, of the additional proteins identified from the broth cultures, 36% in strain A, 35% in strain B-1 and 34% in strain Tra5/5, were not detected in the crude protein extracts, and were only identified after ProteoMiner treatment. These may correspond to lower abundance proteins, and so may also be expressed in the standard plate extracts, but at a concentration below detection levels.

When determining the value of ProteoMiner treatment for the enhanced detection of low abundance proteins, the treated fractions were compared to crude extracts to see how many additional proteins were identified. Previously, the percentage of proteins found after treatment which were not detected in untreated samples has been reported to be 38% for bile extracts (Guerrier *et al*,

2007), and 65% for Urinary protein extracts (Castagna *et al*, 2005), and the number of peaks identified by SELDI-TOF-MS increased from 48 in crude plasma samples, to 330 after treatment with beads (Sihlbom *et al*, 2008). In all these studies, larger numbers of proteins were detected after treatment than before. Conversely, in this study, fewer proteins were identified in the ProteoMiner treated fractions than the crude extracts for each strain. However, there were additional subsets of proteins present in the treated fractions not seen in the crude extracts.

Additional proteins were visible in both the 1D and 2D gel profiles of treated fractions that were not seen in those of the crude broth extracts. Moreover, the additional proteins visualised in the amino-terminal E-bead and carboxy-terminal SE-bead treated fractions appear to be different, indicating that both libraries have roles in the exploration of the so-called deep proteome. This is also true of the proteins identified by LC-MS/MS, where different subsets of proteins were detected in the E-bead and SE-bead treated fractions. In this study, the use of N-terminal E-bead libraries allowed the detection of an additional 65, 88 and 112 proteins not seen in the crude extracts for strains A, B-1 and Tra5/5 respectively, comprising 10%, 17% and 19% of the total number of proteins identified, and the sequential use of the carboxy-terminal SE-bead library identified 83, 38 and 58 additional proteins not seen in any other fraction, comprising 13%, 8% and 10% of total proteins detected in each strain. In total, 23% of proteins found in strain A, 25% of proteins found in strain B-1, and 29% of proteins found in strain Tra5/5 were only detected after ProteoMiner treatment. This is a lower percentage than has been reported previously for biological fluids.

In each strain, the majority of additional protein spots picked from the 2D gel profiles of treated extracts and identified by MALDI-TOF MS had not been seen on the reference maps of the standard plate extracts. This suggests that ProteoMiner treatment has decreased the dynamic range of protein expression, and increased the relative concentration of these lower abundance proteins, so they are now detectable by 2D reference mapping. However, as discussed above, the ProteoMiner treated samples were extracted from broth culture rather than plate culture, so it would be expected that different proteins would be expressed anyway (as seen from the LC-MS/MS protein identifications). The majority of additional proteins had been identified in the standard plate extracts by LC-MS/MS supporting the idea that they were present in the standard plate extracts but at too low a concentration to be detected by MALDI-TOF MS.

The protein extracts eluted from the SE-bead library with carboxylated terminal residues particularly show an increase of protein spots visible in the basic region of the gel compared with both the crude and E-bead treated fractions, as shown clearly by the positions of the red boxes on the SE bead treated gels (*Figure 3.18).* This has also been seen in a number of other biological samples treated with complimentary libraries (Righetti and Boschetti, 2008: Guerrier *et al* 2007). These basic proteins are notoriously difficult to resolve on 2D gels (Antelmann and Hecker, 2010), and so were of particular interest. Therefore a larger pH range (3-10) was used for the 2D gel profiles of these broth samples than was used for the standard plate extracts (pH 4-7). Of the additional proteins identified from the 2D profiles of the ProteoMiner treated samples, five from strain A, seven from strain B-1 and six from strain Tra5/5 had theoretical

pI values outside the 4-7 pH range used for the 2D reference map gels, another reason why they would not have been detected.

In order to determine whether greater numbers of basic proteins were detected in the SE bead treated fraction by LC-MS/MS, theoretical pI values were obtained for each protein identified in the treated fraction of strain A using the ExPASy pI calculator tool (http://expasy.org/tools/pi_tool.html). The same trend was seen in this data, with 31% of the proteins detected in E-bead treated fractions, but not the crude fraction, showing a theoretical pI greater than seven, whereas 41% of the additional proteins detected in the SE-bead treated fraction only, were predicated to be basic with a pI greater than seven.

The additional proteins seen after ProteoMiner treatment included a large proportion of ribosomal proteins, particularly for the 2D MALDI-TOF methodology where 40% additional proteins is strain A, 54% of the additional proteins in strain B-1 and 60% of the additional proteins in strain Tra5/5 were ribosomal proteins. This was less marked for the LC-MS/MS analysis where only 5% of the additional proteins detected after ProteoMiner treatment in strain A, 11% in strain B-1 and 5% in strain Tra5/5 were ribosomal. Ribosomal proteins would not be expected to be of particularly low abundance, but appear to have been much concentrated by treatment with the ProteoMiner libraries, and their detection particularly greatly increased by 2D reference mapping after ProteoMiner treatment. This may be because the ribosomal complex is not fully disrupted during the extraction process, but further treatment with the ProteoMiner libraries disrupts the complex, solubilising individual ribosomal proteins and so allowing them to be visualised. Alternatively, it may be that the

basic nature of the ribosomal proteins makes visualisation hard, as the basic regions of 2D gels are far less easily resolved. The carboxy-terminal SE-bead library may stabilise the visualisation of these proteins on the 2D gel.

Only a small number of additional surface and virulence proteins were identified after ProteoMiner treatment, and the majority of those had also been previously identified in the standard plate extracts. 2D GE followed by MALDI-TOF MS identified one sporulation protein, Spo0A in the treated extracts of all three strains, whereas on the reference maps it was only seen in strains A and Tra5/5. An additional cell surface protein (gi|11525176) was detected in strains B-1 and Tra5/5 but not A. These had both been identified in the standard plate extracts by LC-MS/MS. LC-MS/MS analysis on the broth samples identified additional cell surface, sporulation and flagellar proteins in the treated fractions of each strain which were not identified in the crude extracts. Most of these had been identified in the standard plate extracts, but some were newly detected in each strain.

It is interesting that with the large variation in proteins detected in the standard plate and broth extracts, the surface and virulence proteins expressed remained broadly constant. *Clostridium difficile* is a highly adaptable organism, with the capacity to survive in various habitats both within a host and in the environment. Spore formation plays a key role in allowing this pathogen to persist in hostile environments, but the ability to retain its surface proteins and virulence factors under varying conditions of growth may help it to colonise a wide range of hosts.

It has been reported previously that some proteins species detected in crude samples are not present after ProteoMiner treatment (Thulasiraman *et al*, 2005). This number has been reported to be ~7% of abundant urine proteins (Castagna *et al*, 2005), 5% of red blood cell proteins and 13% of platelets extracted proteins (Righetti and Boschetti, 2008). This is likely to happen if the hexapeptide library does not contain a suitable ligand for a particular protein, preventing it binding to the beads, and hence excluding it from the treated fraction. Alternatively, it may be that the protein – peptide complex is governed by a strong association constant which is incompatible with the eluting agents used. In this study, 28% of the proteins identified in the crude extract of strain A, 30% of proteins in the crude extract from strain B-1 and 34% of proteins from the crude extract of strain Tra5/5 were not detected in the ProteoMiner treated fractions, a higher proportion than has been previously described. It is possible that a larger number of proteins in the hugely complex bacterial extract do not have corresponding ligands within the library, so fewer proteins are captured on the beads. Alternatively, there may be more proteins forming very strong interactions with the bead library, preventing elution.

Recent work by Keidel et al (2010) indicated that the proposed mode of action of the ProteoMiner bead library may be false, and that instead of specific interactions with the hexapeptide library, hydrophobic interactions are the most important forces binding proteins to the ProteoMiner beads. Similar results were seen after treatment with ProteoMiner beads and a number of chromatographic Sepabeads, indicating that the same proteins are bound by both, despite the supposed differences in interaction chemistry (Keidel *et al*, 2010). This may explain why there is still a large range in protein concentrations seen on the 2D gel profiles, and may also help to explain why

such large numbers of proteins were lost from the crude sample after ProteoMiner treatment. If the mechanism of protein binding to the bead libraries is different from that proposed (Thulasiraman *et al*, 2005), then it may be that fewer proteins than expected bind to the library.

4.4 – Western Blotting

Another key factor in the pathogenicity of *Clostridium difficile* and many other bacteria, is the interaction of the organism with the host, and its ability to evade detection by the host immune system. Serum levels of IgA and IgM were found to be significantly higher in asymptomatic carriers of *C. difficile* than symptomatic patients suffering from *Clostridium difficile* associated disease (CDAD) (Mulligan *et al*, 1993), indicating that the immune response of the host plays a key role in preventing illness.

Toxin A is known to be highly immunogenic, and high levels of antibodies in response to toxin A has been shown to significantly reduce risk of recurrent *C. difficile* infection (Kyne *et al*, 2001). This is now the basis of Toxoid vaccines, which induce immune responses against toxins A and B, as treatment in recurrent CDAD (Sougioultzis *et al*, 2005). However, immune responses against a number of other proteins have previously been reported (Pechine *et al*, 2005a); (Ausiello *et al*, 2006); (Drudy *et al*, 2004), and analysis of reactive proteins by western blotting can help to identify proteins that interact with the host, and may be potential virulence factors.

A number of *Clostridium difficile* proteins have been reported to be recognised by the immune system. Antibody levels of Cwp66 (N-terminus), FliC, FliD and a fibronectin-binding protein, Fbp68, were found to be higher in a control group than a group of *C. difficile*-colonised patients, suggesting that a defence mechanism of the host against these proteins could affect the course of infection (Pechine *et al*, 2005a). The exposed part of the flagella has also been shown to be highly immunogenic (Pechine *et al*, 2005b).

During this analysis, a number of proteins from each strain were recognised by the rabbit serum raised against the reference 630 strain. The flagellin protein FliC and the HMW SLP are strongly immunoreactive as has been reported previously (Pechine *et al*, 2005a); (Drudy *et al*, 2004). However, western blotting with human sera was unsuccessful. Two dimensional western blotting of cell wall proteins with human sera has been published, and the S-layer proteins identified as immunoreactive, along with a number of other cell wall proteins (Wright *et al*, 2008). The serum samples used in this analysis had been analysed previously for antibody responses (Sanchez-Hurtado *et al*, 2008), and surprisingly, levels of IgG and IgM for most antigens were similar between the three groups of patients, indicating that there was little difference in the immune responses to infection of the symptomatic and asymptomatic patients. Western blotting was carried out with these serum samples, and the serum pooled from symptomatic cases reacted with more protein bands than the serum from asymptomatic carriers or controls. However these results could not be reproduced here.

During this analysis, one dimensional western blotting with both rabbit and human serum showed an area of strong antigen detection at the gel front (*Figures 3.29, 3.39, 3.40 and 3.41*). This may be due to small fragments of peptides or lipocarbohydrates (lipoteichoic acid-like molecules), which have previously been suggested to be immunoreactive (Sharp and Poxton, 1986). However this was not seen by 2D western blotting.

Western blotting using human serum is often problematic, hence the majority of studies into immunoreactive proteins of *C. difficile* have been done with other methodologies, such as ELISA (Pechine *et al*, 2005b); (Ausiello *et al*, 2006); (Drudy *et al*, 2004). One of the reasons for this is the complexity of the human immune response. Many antibodies to a wide range of antigens are present within the serum, which can mask the reactivity of antigens present at low levels. This may explain the lack of immunoreactive observed here with human serum samples, whereas the use of rabbit serum raised directly against *C. difficile*, which would be expected to contain a much higher concentration of antibodies against this pathogen, identified a number of immunoreactive protein bands.

At least six protein bands reactive to rabbit sera were seen in each strain by 1D western blotting, some of which are likely to correspond to more than one protein of the same size. An additional strongly reactive band of ~66kDa was seen in strains A and Tra5/5 which was absent in strain B-1. 2D western blotting allowed some of these proteins to be identified by matching of reactive profiles to the 2D reference maps.

A number of the immunoreactive proteins were conserved between the three strains, with the flagellin protein, FliC, and the high molecular weight S-layer protein (HMW SLP) clearly present on the reactive profiles of each strain. In addition, a number of reactive spots were present in the centre of the western blot of each strain showing similar mass to the HMW SLP, but a more basic pI. These were matched to a putative amino acid aminotransferase, an elongation factor T, isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase, and glutamate dehydrogenase. Interestingly, the latter is currently used as an *in vitro* diagnostic test for *C. difficile*. The most noticeable difference between the profiles of the three strains was the series of spots in the top right hand corner of the western blot profiles of strains A and Tra5/5, but not strain B-1. This series of spots are likely to correspond to the band at ~66kDa seen in the 1D western blot profiles of the same strains, and was matched to the cell surface protein, gi|115251844, detected in strains A and Tra5/5 but not B-1.

Correct alignment of immunoreactive profile and whole cell extracts was difficult. The use of two Fluorescent dyes which could be visualised together helped this considerably, but there were still problems in matching immunorective proteins to reference maps, particularly those reactive proteins which appear to be of low abundance and cannot be visualised on the whole cell protein extracts. Therefore, many of the identifications of immunoreactive proteins at this stage can only be tentative.

4.5 – Biological Significance

Strains B-1 and Tra5/5 were chosen for this analysis due to the wealth of biochemical data available for these isolates. There is less extensive biochemical

data available for strain A, but the number of studies using 027 ribotype strains is increasing. Previously reported virulence factors for strains B-1 and Tra5/5 were examined in greater detail, and compared to the proteins detected in each strain (*Table 4.1)*.

***Table 4.1** – Previously described biochemical data for reported virulence factors of C. difficile and the corresponding proteins detected*[1]. (Borriello *et al*, 1987); (Krishna *et al*, 1996) (Borriello *et al*, 1988b; Seddon *et al*, 1990; Tasteyre *et al*, 2000a) (Davies and Borriello, 1990) (Borriello *et al*, 1988a) (Stabler *et al*, 2009; Warny *et al*, 2005)[2]

	Strain B-1		**Strain Tra5/5**		**Strain A**	
	Bio. data	**Proteins**	**Bio. data**	**Proteins**	**Bio. data**	**Proteins**
Virulence	high	-	low	-	High	-
Toxin A	>195 ng/ml	+	~25ng/ml	+	848μg/L	+
Toxin B	1:40,960	+	1:40,960	-	180μg/L	+
Mucosal adherance	+	HMW SLP - - other SLPs	n/a	HMW SLP LMW SLP SLP13[3] other SLPs	n/a	HMW SLP LMW SLP SLP13[3] other SLPs
Fimbrea	-	-	+	Pilin *	n/a	-
Flagella	+	FliC Flagella hook	n/a	FliC -	n/a	FliC Flagella hook
Motility	+	2 chemotaxis proteins	n/a	-	+	1 chemotaxis protein
Capsule	+	-	+	-	n/a	-
Hydrolytic enzymes - chondroitin-4-sulphatase - hyaluronidase - collagenase -heparinase	 + + weak +	None detected	 + + weak +	None detected	n/a	None detected

[1] *For full lists of detected proteins see Appendix Tables A1and A2*

[2] *For full references see section 2.1.2.*

[3] *Table 3.2. Corresponds to spot J12, Figure 3.8 and Figure 3.10.*

* *Detected in broth samples only*

The virulence factors described in Table 4.1 are discussed in more detail below, along with some other previously reported virulence factors.

4.5.1 - Toxins

All three strains examined in this study showed an intact pathogenicity locus by genome sequencing (personal communication), so would be expected to produce toxin A and toxin B (Hammond and Johnson, 1995).

Both toxins A and B were identified by 1D GE LC-MS/MS analysis, but not on the 2D reference maps. This is likely to be due to the large size of these proteins (270 and 308 kDa) making them difficult to resolve by gel electrophoresis. Similar problems with visualising the toxins on gels were reported previously by Mukherjee *et al* (2002). By LC-MS/MS, toxin A was identified in all strains, but interestingly, toxin B was not detected in strain Tra5/5. This may indicate that strain Tra5/5 produces lower levels of toxin B than the other two strains, although strains Tra5/5 and B-1 have been shown previously to produce comparable amounts of toxin B in vitro (Borriello *et al*, 1987). The same variation was seen in both the standard plate extracts and the broth extracts used for ProteoMiner treatment, with toxin A being present in the broth extracts of all three strains, while toxin B was pesent only in strains A and B-1 but not in strain Tra5/5. No binary toxin was detected in any of the strains under any growth conditions.

Toxins A and B are produced in response to stress, including temperature (Karlsson *et al*, 2003). Toxin production has been shown to be growth dependent (Hundsberger *et al*, 1997) with levels of transcripts of both toxin A and toxin B staying low during logarithmic growth, but high during stationary phase. Toxin production is also affected by available nutrients and has been shown to be repressed by glucose (Dupuy and Sonenshein, 1998), induced by biotin limitation (Yamakawa *et al*, 1998) and varies according to amino acid composition of the growth medium (Haslam *et al*, 1986). All three strains have been grown under the same conditions, so lack of detectable toxin B in Tra5/5 cannot be attributed to differences in growth conditions. Despite being grown on nutrient rich Colombia blood agar, all strains produced toxin A.

The role toxin B plays in the virulence of *C. difficile* is still disputed (Lyras *et al*, 2009; Kuehne *et al*, 2010), but it has been shown to disrupt epithelial integrity in the host cells (Hecht *et al*, 1992) and function as a potent enterotoxin in its own right (Savidge *et al*, 2003). If, under both sets of growth conditions used here, strain Tra5/5 does not produce toxin B, despite the in tact pathogenicity locus, it may go some way to explaining the observed lower virulence of this strain in the hamster model. However, as Tra5/5 has been shown to produce comparable levels of toxin B to strain B-1 under some *in vitro* conditions, it would be interesting to see how toxin levels between these two strains compared under other growth conditions.

It has previously been reported that the 027 ribotype strains produce higher levels of toxin than less epidemic strains due to mutation in the *tcdC* negative regulator. Although both toxins A and B were detected in the 027 ribotype

strain A and the less epidemic strain B-1, no quantitative information can be gleaned from the LC-MS/MS analysis performed here, so expression levels of toxin between strains cannot be compared.

4.5.2 – Sporulation

Clostridial species, like bacilli, have the ability to form endospores in response to hostile environments as a means to survive, so can be considered to have two distinct lifecycle phases with a very different set of ORFs expressed in each. The mechanism of spore formation and the spore structure of *Bacillus subtilis* have been well characterised, but the equivalent mechanisms in clostridia are poorly understood.

In *Bacillus subtilis*, DNA-macroarray analysis suggested that 61% of predicted ORFs are significantly expressed in vegetative cells (Eymann *et al*, 2004).

The genome of the 630 strain of *C. difficile* contains 60 predicted sporulation/germination proteins (Sebaihia *et al*, 2006), and the protein composition of spores and vegetative cells have been shown to be very different, with ~336 proteins identified as components of purified spores (Lawley *et al*, 2009). The standard protein profiles here resemble the vegetative proteome rather than the proteome of purified spores (Lawley *et al*, 2009) but it is likely that a number of spores are present in the cultures used for extract. A number of proteins present in extracts of purified spores were also detected in the standard extracts during this analysis. Stress response proteins were detected, including rubrerythrins and tellurium resistance proteins, found in all strains, and a putative catalase, found only in strain A. Ribosomal components, chaperones and proteases have also been shown to be abundant in spores.

Fewer chaperones were identified here than in the spore extracts, four compared to 16, but a large number of ribosomal components were detected, particularly after ProteoMiner treatment.

The presence of these proteins in the standard extracts may indicate that spores present in the standard plate cultures have been disrupted by the extraction process and spore proteins are present in the extract. Alternatively, these proteins could be ubiquitously expressed during both vegetative growth and in spores. Infection of a host is stressful for the pathogen, so it is likely that a number of stress response proteins are also produced during vegetative growth.

Spo0A is an essential factor in sporulation initiation in both *B. subtilis* (Piggot and Hilbert, 2004) and *C. difficile,* where it has also been shown to be linked to toxin production in *C. difficile* (Underwood *et al*, 2009). In this analysis, it was found in all three strains in both broth and plate extracts. The other sporulation protein identified in all strains in both the plate and broth samples was an SpoIIIJ associated protein. SpoIIIJ activates the transcription factor sigma G, which is required for sporulation, but it has been shown that SpoIIIJ is also ubiquitously expressed in vegetative cells (Serrano *et al*, 2003).
In bacillus, the sigma F regulator becomes active immediately after asymmetric cell division in the daughter cell destined to become the prespore (Piggot and Hilbert, 2004), a stage of sporulation termed stage II. Until division, this protein is held inactivated by the anti-sigma F factor, which was detected in all three strains, indicating the majority of cells within the culture were not yet sporulating. One additional putative stage II sporulation protein was found in strain A.

Interestingly, an Spo0B associated protein was detected in each of the three strains, but the Spo0B protein from bacilli does not appear to have any homologues within the *C. difficile* genome. Instead, spo0A appears to be directly phosphorylated by orphan kinases (Underwood *et al*, 2009) only one putative sensor histidine kinase was detected in this analysis 630_gi|115251546 and this was only found in strain A.

A number of latter stage sporulation proteins stage IV and V were detected in different strains.

In bacillus, the mature spore consists of a core, housing the chromosomes and surrounded by a lipid membrane. This is then surrounded by the cortex, a thick layer of peptidoglycan followed by a multilayered structure called the coat (Driks, 1999). Two spore structural proteins were identified, but only in strain A, a putative spore coat protein, and a Spore Cortex protein. This may indicate that there are sporulation differences and that more cells within the strain A culture have progressed along the sporulation pathway. Alternatively, strain A may produce these structural spore proteins in greater abundance, or at different times to the other two strains. Emerging 027 ribotype strains have been reported to show increased sporulation rates compared to non-epidemic 027 strains and those of other ribotypes (Akerlund *et al*, 2008), supporting this hypothesis.

4.5.3 - S-layers

S-layers have been identified in hundreds of different species from all areas of the bacterial kingdom, and have been shown to be essential for the virulence of some bacterial pathogens such as *Aeromonas salmonicida* and *Campylobacter fetus* (Sara and Sleytr, 2000). The S-layer of *Clostridium difficile* is unusual, in that it consists of two subunits forming a lattice structure, whereas the majority of bacterial S-layers consist only of one (Sara and Sleytr, 2000). Most *Clostridium difficile* strains express a high molecular weight (HMW) protein and a low molecular weight (LMW) protein, which are derived from a single gene, *slpA* (Calabi *et al*, 2001) and cleaved post-translationally. The protein precursor contains a typical signal sequence indicating that it is translocated across the membrane via the general secretory pathway before cleavage occurs (Calabi *et al*, 2001).

During this analysis, the SlpA protein was identified in all strains, and showed considerable amino acid variation between strains, and only around 50% homology between strains by BLAST.. The data from the reference maps and DIGE analysis suggests that LMW-SLP is not present in the strain B-1 extract, indicating that the cleavage event could be different in strain B-1 to the other two strains. A similar variation in the SLP protein has been seen by Calabi and Fairweather who identified a strain, designated 167, which showed an unusual S-layer 1D profile (Calabi and Fairweather, 2002). A predominant protein of 39kDa representing the HMW SLP was detected, but only minimal amounts of the predicted, 20kDa LMW protein was seen. The un-cleaved pre-protein of this strain had a predicted molecular mass of 62,312 Da, lower than that of other isolates, which is similar to the predicted mass for strain B-1 (64,703Da). The

authors proposed that the lack of LMW SLP could be due to rapid degradation of this subunit, and showed that this did not affect the pathogenicity or cellular viability of the strain. A different S-layer lattice system may be employed in this, and the B-1 strain.

The HMW SLP shows homology to the Bacillus family of autolysins (*CwlB* and *CwbA)* with N-acetyl muramoyl-L-alanine amidase activity (Karjalainen *et al*, 2001) particularly in the predicted cell wall binding domain. HMW SLP has also been shown to have amidase activity (Calabi *et al*, 2001). Multiple homologues to the *SlpA* gene encoding the S-layer protein have been found within the genome which show similarity in the *CwlB*-like domain, and may have similar surface functions.

Calabi *et al* (2001) identified 29 ORFs encoding proteins showing homology to this autolysin. Many of these ORFs were clustered near *SlpA*, and others were located throughout the genome. In-depth analysis of six of these ORFs showed them to be expressed in growing cultures by RT-PCR. In *Campylobacter fetus* several genetic SLP cassettes have been identified (Dworkin and Blaser, 1997), but only a single gene is expressed at a given time, whereas the paralogues in *C. difficile* were found to be expressed concurrently (Calabi *et al*, 2001).

It has been reported that the *SlpA* gene is located on a locus containing 17 open reading frames, 11 of which encode proteins containing domains homologous to the postulated cell wall anchoring domain of *SlpA*, (Karjalainen *et al*, 2001). The structure and organisation of this locus suggests that these genes could be co-transcribed and form an operon (Savariau-Lacomme *et al*, 2003), and the

similarity of the *SlpA* homologues to the cell wall anchoring domain of the *B. subtilis* autolysin (Karjalainen *et al*, 2001) suggests modular evolution and gene duplication from a common ancestor of *B. subtilis* and *C. difficile.*

Well characterised components of this genetic locus include Cwp66 (gi|115251842) and Cwp84 (gi|115251840). Cwp66 has been shown to act as an adhesion in heat-shocked cells (Waligora *et al*, 2001) and is thought to undergo similar cleavage events to SlpA. Cwp84 is located immediately downstream of Cwp66 and encodes a putative cysteine protease. It has been suggested that Cwp84 plays a role in maturation of SlpA into the HMW and LMW SLPs, as genetic knockout mutants expressed only the uncleaved SlpA precursor. However, it was shown not to directly play a critical role in disease as the mutant strains were still capable of causing disease in the hamster model (Kirby *et al*, 2009)

The *SlpA* locus also contains the *SecA* translocase gene, encoding an essential component of the general secretory pathway, the route by which SlpA is thought to be exported. The two genes are adjacent to one another, indicating that the two may be transcribed together (Calabi *et al*, 2001).

In this analysis, 10 surface proteins from this genetic locus were detected, including SlpA, Cwp66 and Cwp84. It has been reported previously that Cwp66 and Cwp84 are transcribed only in the early exponential phase, but in this study they were identified in the standard extracts grown for 64 hours in all three strains, when the majority on cells are in stationary phase. The SecA translocase was also identified in all strains.

This genetic locus was examined in more detail, and the genomic information from strains A, B-1 and Tra5/5 was compared to that of strain 630 (*Figure 4.1)* using the progressive Mauve algorithm (Darling *et al*, 2010).

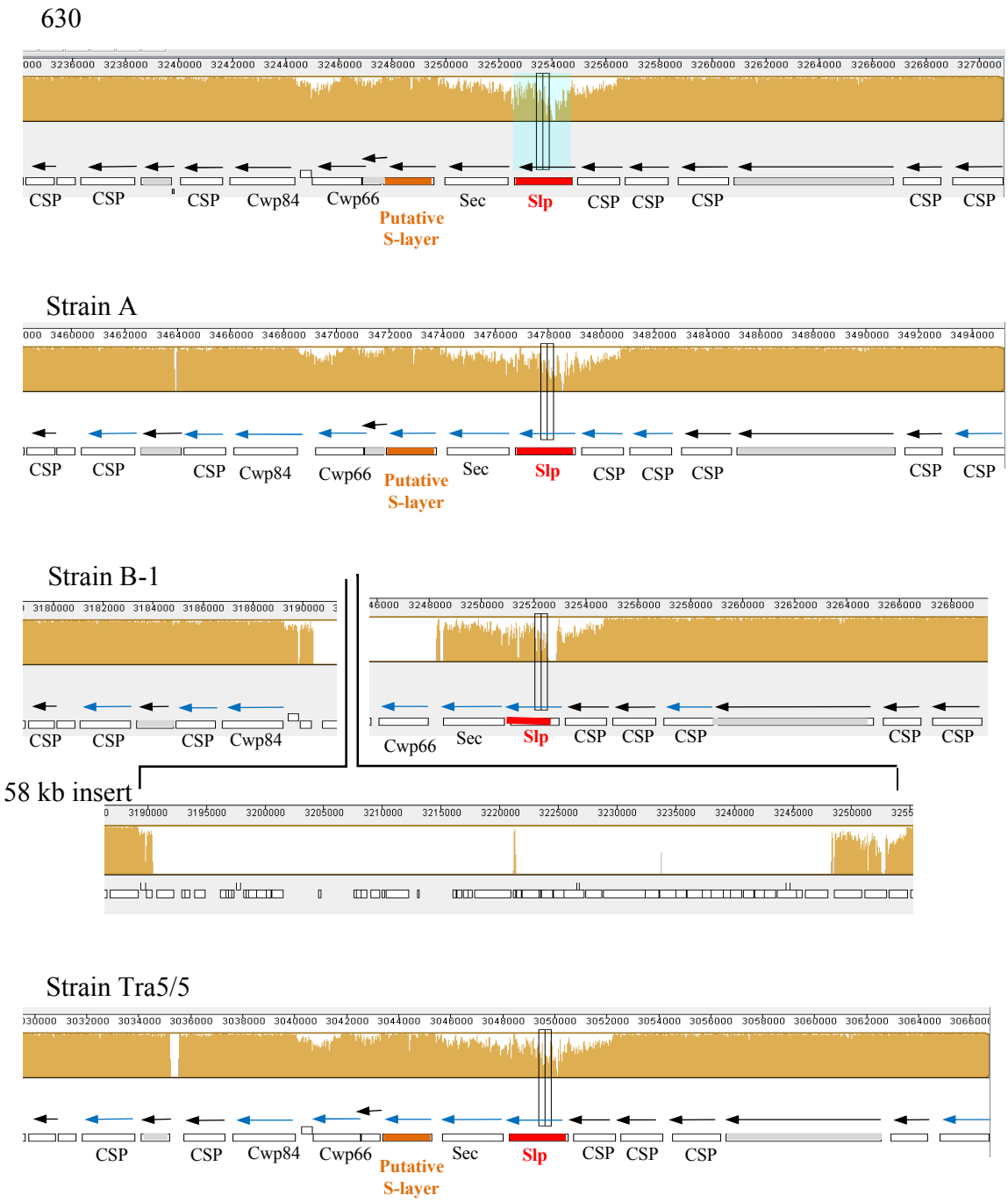


Figure 4.1– *Organisation of open reading frames in the SlpA operon for strains A, B-1, Tra5/5 and the 630 reference strain. The 37kb fragment described by Karjalainen et al (2001) was looked at in detail and the 17 ORFs encoded within this fragment are shown above for strain 630, and matched to other the three strains. Yellow bars represent*

areas of homology between the strains, with the height of the bar representing the level of homology. Genes containing more variation, such as the SlpA gene, show white areas within the yellow. The SlpA gene is shown in red in all strains, and highlighted in green in strain 630, the gene encoding the putative S-layer protein (gi|115251844) is shown in orange. Hypothetical or putative ORFs are shown in grey. Arrows above the genes show the direction of transcription, and where the arrow is shown in blue, the corresponding protein was detected in this analysis. CSP notates genes encoding cell surface proteins (Slp homologues). The genetic fragment for strain B-1 shows considerable differences to the other three genomes, with an insertion of approximately 58 Kb, in the position shown by the black lines bisecting this genome above. This insertion contains 50 ORFs, the majority of them putative and uncharacterised.

The majority of ORFs in this genetic locus were found in all strains, with the exception of a small hypothetical ORF (CD2790), encoding a hypothetical protein of 235aa, and the adjacent ORF (CD2791) encoding the 623aa putative S-layer protein (gi|115251844), which, as discussed above (Table 3.2), was detected in strains A and Tra5/5 but not B-1. This protein corresponds to the reference map spot J12 which was detected in strains A and Tra5/5 but not B-1 by both MALDI-TOF MS and LC-MS/MS, and shown to be upregulated in these strains compared to strain B-1 by DIGE.

These ORFs appear to have been deleted from this locus in strain B-1. In addition, in strain B-1, this genetic locus includes an insertion of approximately 58Kb between the ORFs of Cwp66 and Cwp84, containing 50 ORFs, most of which are putative or uncharacterised. However, this does not seem to block the transcription of these ORFs, as many of them, including Cwp66, Cwp84, SecA and SlpA, were confirmed by the detection of the corresponding proteins.

Therefore, genetically, as well as proteomically, strain B-1 appears to show a markedly different profile to the other two strains with regard to S-layer and surface proteins.

The S-layer proteins from *Clostridium difficile* have been shown to induce an inflammatory response (Ausiello *et al*, 2006) and detectable levels of anti-SLP antibodies have been found in human sera (Drudy *et al*, 2004). Higher levels of IgM antibodies were found in asymptomatic patients than in those that went on to develop CDAD, and patients suffering from recurrent CDAD had significantly lower IgM levels than those suffering a single episode of infection, suggesting that strain specific antibody response to SLPs might be important in clinical outcome of infection.

The HMW protein is fairly conserved among strains, but the LMW protein shows considerable variation between different *C. difficile* isolates (McCoubrey and Poxton, 2001); (Calabi *et al*, 2001). Cross immunoreactivity was seen between the HMW proteins, but not the LMW proteins (Calabi *et al*, 2001). In this study, the HMW protein is strongly immunoreactive, but the spots corresponding to the LMW protein in strains A and Tra5/5 (figures) do not react with the serum primary antibody, indicating that there is a high divergence from the corresponding 630 protein against which the rabbit serum was raised. The divergence of this LMW SLP may help the pathogen evade host immune systems.

SLP homologues have also been shown to trigger immune responses, so the immunoreactive spots were compared to the theoretical mass and pI values of these proteins. Despite high interstrain variability of the *Cwp66* gene, antibodies

against this protein recognised Cwp66 from different isolates. Cwp66 has a predicted molecular weight of 67kDa, so could correspond to the band at ~65kDa on the 1D Western blot, and the series of spots on the 2D Western blot that were present in strains A and Tra5/5 but not B-1. This protein was detected in all strains, but as it showed homology between strains A and Tra5/5, and much lower homology to strain B-1, it is possible that Cwp66 from B-1 is different enough to escape detection by the Cwp66 antibody raised against the 630 protein. It is possible that this series of spots corresponds to a number of different homologues, all of which exhibit variability in B-1 preventing cross reactivity.

Cwp84 is highly conserved and has been shown to be strongly immunoreactive (Pechine *et al*, 2005b). Cwp84 has a predicted molecular weight of 87kDa. There was a faintly immunoreactive band visible in the 1D Western blot with rabbit serum that appeared to be present in strain Tra5/5 but not the other two strains which may correspond to this protein. However, as it was detected in all three strains, and was highly conserved, it would be expected that if this band corresponded to Cwp84, it would be present in all strains.

Alternatively, the series of immunoreactive spots seen in strains A and Tra5/5 but not B-1 could correspond to the additional S-layer protein gi|115251884, which has been discussed above and is present in the genomes of these strains but not that of B-1. This protein could undergo a modification event, causing it to appear as a series of spots. It was identified as a single spot on the reference maps, but it is possible that additional spots were obscured by other, non immunoreactive, proteins, so can only be seen on the reactive profile.

This analysis indicates that the S-layer proteins, and SLP homologues are similar in amino acid composition and reactivity in the hypervirulent strain A and the low virulent Tra5/5, but that strain B-1 exhibits a markedly different surface protein profile. B-1 shows different expression and modification of the SlpA protein, in that the LMW SLP was not detected in this strain. In addition it exhibits genetic variation, with differences in the organisation of the *SlpA* locus in this strain, including the absence of the gene encoding the surface protein gi|115251844, and an adjacent small hypothetical ORF. The fact that the B-1 strain does not appear to express the LMW SLP, and shows a different genetic S-layer profile, supports previous conclusions suggesting that although the S-layer is likely to be important in adhesion and interaction with the host, many other factors must be involved in virulence determination.

Four cell surface proteins were detected in the virulent strains A and B-1, but not in the low virulence Tra5/5 strain. Only one of the genes encoding these proteins was located in close proximity to *SlpA* and none of them have been previously described in detail in the literature. One protein (SLP11, gi|115251237) was identified as a cell wall hydrolase with homology to the CwlD N-acetylmuramoyl-L-alanine amidase protein family involved in degradation of peptidoglycan. It has been suggested that in bacillus, this protein plays a role in the production of spore-specific peptidoglycans (Gilmore *et al*, 2004), so may play a role during sporulation. Two proteins, gi|115249861 (313aa) and gi|115251839 (525aa) contain three putative cell wall binding motifs (CW-binding repeat 2), which are found in multiple copies in surface proteins including amidase enhancers and adhesions, as well as LytB-like cell wall

binding domains. The presence of these domains has led to their classification of putative surface proteins, and indicates that these proteins may play a role in adhesion. The fourth protein, gi|115251764, is slightly larger (656aa) and as well as containing three copies of the CW-binding repeat 2 and two LytB cell wall binding motifs, this protein contains an L,D-transpeptidase catalytic domain of the YkuD superfamily. This family of proteins can act as transpeptidases, giving rise to an alternative pathway for peptidoglycan cross-linking. This can give bacteria resistance to beta-lactam antibiotics, as the beta-lactam sensitive penicillin binding proteins that usually catalyse the final peptidoglycan cross linking stages of cell wall synthesis, are no longer required (Biarrotte-Sorin *et al*, 2006).

If these proteins are indeed not produced in the lower virulence Tra5/5 strain, or produced at reduced levels compared to the more virulent B-1 and A strains, this could indicate that there is a difference in the adhesive properties, or peptidoglycan crosslinking and degradation properties of this strain. This may even affect the antibiotic resistance profile of Tra5/5. These differences may play a role in the reduced virulence of Tra5/5 compared to strains B-1 and A.

4.5.4 – Flagella and motility

Flagella are cell surface structures found on many bacterial species. The structure is made up of three main parts, a basal body, functioning as a motor, a torsion hook and a hollow filament making up the flagellum tail, and flagella have been shown to contribute to pathogenesis in a range of bacteria in a number of ways (Ramos *et al*, 2004).

Firstly, flagella confer motility, allowing access of the pathogen to mucosal tissues. Secondly, flagella can act as an adhesin. Crude preparations of *Clostridium difficile* flagella have been shown to associate with the cecal mucus of mice, and non-flagellate strains of *C. difficile* associate with cecal tissue at a rate that is 10-fold lower than that of flagellate strains (Tasteyre *et al*, 2001). Finally, it has been suggested that flagella can play a role in sensing of environmental conditions, such as 'wetness' (Wang *et al*, 2005) and hence may control expression of genes and virulence factors under certain conditions (Boin *et al*, 2004).

Flagellin or FliC, the main protein component of the fillment, was identified as a series of different spots on each 2D gel, indicating post-translational modification. However, DIGE analysis showed different protein migration between the three strains. The amino acid sequence of flagellin is highly conserved between the three strains, and similar to that of strain 630, as deduced from genome sequencing (personal communication). FliC from strains A and Tra5/5 shows a 29aa insertion at the sequence start, which is not present in strains B-1 or 630. The flagellin of strains B-1 and Tra5/5 is otherwise identical, and shows two amino acid differences to the flagellin protein from strain A, with a serine at position 188 instead of threonine, and asparagine at position 227 instead of aspartic acid. This high level of homology means that the difference in flagellin migration is unlikely to be due to amino acid differences, suggesting that variation in the modification process between strains causes the difference in spot position. Similar changes in flagellin protein migration have been seen in *Campylobacter jejuni*, with proteins from a putative flagellar

glycosyltransferase knock out mutant migrating at a less acidic pI within the gel than those with an active glycosyltransferase (van Alphen *et al*, 2008).

Glycosylation of flagella has been shown to be important in many bacterial species, affecting motility, adhesion and pathogenesis (Logan, 2006). It was recently shown that the flagellin of *Clostridium diffile* is glycosylated in the central, surface exposed region, and that glycosylation is required for filament assembly and subsequent motility (Twine *et al*, 2009). Variation in the glycans of different strains was observed, indicating that glycosylation differences are likely to be responsible for the migration differences observed.

Differences in the flagellar-associated glycosytransferases have been shown between epidemic O27 ribotypes and non-epidemic strains (Stabler *et al*, 2009), again indicating that the glycosylation process may vary between isolates, and that this may have an affect on virulence. Differences in the adjacent genes to the flagellar-associated gene loci between the epidemic O27 ribotypes and the reference 630 strain have been observed to include an additional glycosyl transferase. These glycosylation differences may be responsible for the observed difference in motility, with the O27 ribotype strains showing greater motility than the 630 reference strain (Stabler *et al*, 2009). A difference in modification (most likely glycosylation) of FliC between strains A and B-1 can be clearly seen by the differential migration shown by DIGE and, given the previous observations in other epidemic 027 ribotype, may contribute to the enhanced virulence of this strain.

Flagella proteins have been implicated in attachment of *C. difficile* to the mucus layer of the intestine, as well as the caecal wall (Tasteyre *et al*, 2001) and so are considered to enhance *C. difficile* virulence. If, as seems probable, the difference in flagellin migration is due to differential glycosylation, this could affect structure and function of the flagella, altering the role played in motility and adhesion. This could contribute to the differences in virulence between the different strains. In addition, the flagellar hook protein was not identified in strain Tra5/5, indicating the structure of the flagella basal body in the low-virulence strain Tra5/5 could also be different. A number of other flagellar proteins notably flagellar motor switch proteins, were detected in strain B-1 only, again suggesting difference in the flagellar structure of the three strains.

Comparative genomic assays have indicated that flagella production and motility may not be a common feature for *Clostridium difficile* isolates, and that flagellar gene loci are not necessarily conserved among virulent strains (Stabler *et al*, 2006). It was reported previously, by RT-PCR, that FliC is expressed in both flagellate and non-flagellate strains (Tasteyre *et al*, 2000b), so it is possible that strain Tra5/5 is non-flagellate, despite the detection of the FliC protein. In addition, FliC was not found to be up-regulated in strain Tra5/5 by DIGE, whereas it was for the other two strains. Data on the motility and flagella production of this strain has not previously been reported, but if this strain is indeed non-flagellate, for example due to differences in the basal body and hook, it could go some way in explaining the low virulence reported for this strain. The presence or absence of flagella in this strain could be verified by Electron Microscopy.

Another component to bacterial motility is chemotaxis, which regulates the movement of bacteria towards an attractant or away from a repellent. It has previously been shown that gut mucus of animals and humans acts as a chemoattractant for *C. difficile* (Boriello and Bhatt R, 1995)}, making chemotaxis proteins potential virulence factors for this organism, as has been shown for other organisms such as *Vibrio cholerae* (Boin *et al*, 2004). Two chemotaxis proteins were identified by this analysis, but neither was detected in strain Tra5/5, again suggesting motility of this, less pathogenic, strain may be different.

4.5.5 – Additional proteins of interest

Hydrolytic enzyme production has previously been linked to virulence in *C. difficile* (Seddon *et al*, 1990) and strains B-1 and Tra5/5 have been shown to produce chondroitin-4-sulphatase, hyaluronidase, heparinase and collagenase. However none of these enzymes were detected during this analysis. Similarly, proteins making up the capsule, which has also been linked to virulence, and shown to be present in strains B-1 and Tra5/5 (Davies and Borriello, 1990), were not detected here.

Fimbriae have been reported in *C. difficile*, although their role in virulence is disputed (Borriello, 1998). Here, the pilin protein making up fimbriae was detected in the low virulence strain Tra5/5 only, and in the protein extracts from broth, not the standard plate extracts. This suggests that under the standard conditions used here, none of the three strains are expressing fimbriae to a detectable level.

Infection of a host by a bacterial pathogen causes stress to both the host and invading organism, and it has been proposed that this stress induces chaperones and heat shock proteins which may play a role in infection. The roles of chaperones in bacterial infection has been reviewed (Henderson *et al*, 2006) and this class of proteins have been shown to be involved in adhesion for a number of organisms. Evidence for the role of the chaperone Hsp60 as an adhesion in *C. dificile* has been presented (Hennequin *et al*, 2001).

This analysis identified three heat shock proteins in all three strains, the 60 kDa chaperonin (Hsp60 630_gi|115249204), Hsp90 630_gi|115249282 and a putative GrpE 630_gi|115251516. Hsp90 and GrpE were also identified in all strains in broth, whereas Hsp60 was only found in strain A in broth extracts.

Another class of protein found to be up-regulated in the virulent A and B-1 strains but not in the less virulent Tra5/5 strain are the ABC transporter substrate binding proteins. Such proteins have been shown to function as adhesins, and potentially linked to virulence in a number of bacterial species including *Campylobacter jejuni* (Leon-Kempis *et al*, 2006) and *Neisseria meningitidis* (Li *et al*, 2009). Again, these proteins may play a role in attachment of the bacteria to the host cells during an infection. Interestingly, more transport related proteins were found in strain A than in strains B-1 and Tra5/5 (31, 27 and 17 respectively). Eight ABC transporter proteins were identified in the A strain only, but fewer were identified uniquely in strains B-1 and Tra5/5, four and two respectively. This may give rise to different adhesive properties between the strains.

In depth examination of virulence factors has identified a number of strain to strain differences which could contribute to the varying virulence of the three isolates discussed here. The association between *C. difficile* and CDAD was first demonstrated over 30 years ago by Larson *et al* (1978)(Larson *et al*, 1978), and by analogy with other toxin producing species, and the conclusion was reached that the toxins are the sole determinates of *C. difficile* pathogenicity. Strain B-1 used here is derived from this study, and this, and other strains, were subsequently used in a number of studies in which this view was reinforced. Both animal models and tissue culture methods were devised to assess virulence, and some 30 years later, these remain the most widely accepted standards. Although a toxin detection kit is now available, its sensitivity in direct faecal samples is poor, and it does not differentiate between toxins A and B. However, the association between toxin production and pathogenicity profiles of *Clostridium difficile* is still not fully understood, and a plethora of virulence factors are likely to mediate host colonisation, adhesion, proliferation and disease progression.

Apart from the S-layer proteins, most of these factors remain poorly characterised, largely due to the lack of appropriate techniques. Proteomics offers new approaches to detect toxins directly and opportunities to study other potential virulence factors across a broad range of isolates, to asses their relative contribution.

The data presented here confirms that the pathogenicity of this organism is not a simple process, and that many potential virulence factors, although not required to cause disease, may contribute to the process. It addition, roles of the

commensal gut flora and the host immune response, not discussed in detail here, are likely to be crucial in determining the outcome of infection. As this organism continues to cause outbreaks in hospitals around the world, placing additional burden on healthcare facilities, ongoing research in this area, and the continuing attempt to elucidate the virulence determinants of *C. difficile* will continue for the foreseeable future. Proteomics should be included as an essential component of these programmes.

Conclusion

Clostridium difficile is a very diverse species, as has been widely reported in the literature. Recent advances in sequencing technology have accelerated the acquisition of a wealth of genomic data for this organism and has provided further evidence of the large genetic variation between different strains (He *et al*, 2010; Sebaihia *et al*, 2006; Stabler *et al*, 2006; Stabler *et al*, 2009). Full genome analysis has shown a large number of mobile genetic elements and a high rate of mutation and gene acquisition from a range of organisms from the commensal gastrointestinal flora such as *Bacteroides*, *Enterococcus, Fusobacterium* and *Ruminococcus* (DBHT, manuscript in preparation).

There has been a changing pattern of epidemiology in circulating strains (Freeman *et al*, 2010), with notable changes in the antibiotic resistance profiles of newly emerging isolates. Many of these differences will be manifested, and hence more readily apparent, in the proteome. Prior to this study, no in-depth proteomic analysis of this organism had been undertaken.

The proteomic tools used for these analyses are highly technically challenging, laborious, time consuming and costly, and therefore, the number of strains that can be included in such a study is limited. New nanoscale methods are developing for rapid, direct analysis. This will facilitate a wider coverage of strains. The recent shift in predominant strains from the 027 ribotype increasingly towards the 078 ribotype can be studied to help elucidate the selective pressures that drive the transition within genetic populations and lead

to new virulent clonal types. The present study will provide a sound scientific framework from which further studies may be developed.

References

1. **Akerlund, T., Persson, I., Unemo, M., Noren, T., Svenungsson, B., Wullt, M., and Burman, L. G.** 2008. Increased sporulation rate of epidemic *Clostridium difficile* type 027/NAP1. *Journal of Clinical Microbiology*, **46**, 1530-1533.

2. **Alam, S. I., Bansod, S., Kumar, R. B., Sengupta, N., and Singh, L.** 2009. Differential proteomic analysis of *Clostridium perfringens* ATCC13124; identification of dominant, surface and structure associated proteins. *Bmc Microbiology*, **9**,

3. **Alfa, M. J., Kabani, A., Lyerly, D., Moncrief, S., Neville, L. M., Al Barrak, A., Harding, G. K. H., Dyck, B., Olekson, K., and Embil, J. M.** 2000. Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile* responsible for a nosocomial outbreak of *Clostridium difficile*-associated diarrhea. *Journal of Clinical Microbiology*, **38**, 2706-2714.

4. **Anonymous**. 16-6-2005. Outbreak of *Clostridium difficile* infection in a hospital in south east England. *CDR Weekly*, **15**,

5. **Antelmann, H. and Hecker, M.** 2010. The Proteomic Road Map to Explore Novel Mechanisms of Bacterial Physiology. In *Mass Spectrometry for Microbial Proteomics* (Shah H. N., Gharbia S. E. Eds), p.339. John Wiley and Sons Ltd, West Sussex.

6. **Ausiello, C. M., Cerquetti, M., Fedele, G., Spensieri, F., Palazzo, R., Nasso, M., Frezza, S., and Mastrantonio, P.** 2006. Surface layer proteins from *Clostridium difficile* induce inflammatory and regulatory cytokines in human monocytes and dendritic cells. *Microbes and Infection*, **8**, 2640-2646.

7. **Barbut, F., Decre, D., Lalande, V., Burghoffer, B., Noussair, L., Gigandon, A., Espinasse, F., Raskine, L., Robert, J., Mangeol, A., Branger, C., and Petit, J. C.** 2005. Clinical features of *Clostridium difficile*-associated diarrhoea due to binary toxin (actin-specific ADP-ribosyltransferase)-producing strains. *J.Med.Microbiol.*, **54**, 181-185.

8. **Bartlett, J. G., Moon, N., Chang, T. W., Taylor, N., and Onderdonk, A. B.** 1978. Role of *Clostridium difficile* in Antibiotic-Associated

Pseudomembranous Colitis
2. *Gastroenterology*, **75**, 778-782.

9. **Bartlett, J. G., Taylor, N. S., Chang, T. W., and Dzink, J. A.** 1980. Clinical and Laboratory Observations in *Clostridium difficile* Colitis. *American Journal of Clinical Nutrition*, **33**, 2521-2526.

10. **Bernardini, G., Braconi, D., and Santucci, A.** 2007a. The analysis of *Neisseria meningitidis* proteomes: Reference maps and their applications. *Proteomics*, **7**, 2933-2946.

11. **Bernardini, G., Renzone, G., Comanducci, M., Mini, R., Arena, S., D'Ambrosio, C., Bambini, S., Trabalzini, L., Grandi, G., Martelli, P., Achtman, M., Scaloni, A., Ratti, G., and Santucci, A.** 2004. Proteome analysis of *Neisseria meningitidis* serogroup A. *Proteomics*, **4**, 2893-2926.

12. **Bernardini, Giulia, Arena, Simona, Braconi, Daniela, Scaloni, Andrea, and Santucci, Annalisa**. 2007b. Novel identification of expressed genes and functional classification of hypothetical proteins from *Neisseria meningitildis* serogroup A. *Proteomics*, **7**, 3342-3347.

13. **Biarrotte-Sorin, S., Hugonnet, J. E., Delfosse, V., Mainardi, J. L., Gutmann, L., Arthur, M., and Mayer, C.** 2006. Crystal structure of a novel beta-lactam-insensitive peptidoglycan transpeptidase. *Journal of Molecular Biology*, **359**, 533-538.

14. **Boin, M. A., Austin, M. J., and Hase, C. C.** 2004. Chemotaxis in *Vibrio cholerae*. *Fems Microbiology Letters*, **239**, 1-8.

15. **Boriello, S. P. and Bhatt R**. 1995. Chemotaxis by *Clostridium difficile*. In *Medical and Dental Aspects of Anaerobes* (Duerden B. I., Wade J. G., Brazier J. S., Eley A., Wren B. and Hudson M. J., Eds), p. 241. Science Reviews Ltd, Middlesex.

16. **Borriello, S. P.** 1998. Pathogenesis of *Clostridium difficile* infection. *J.Antimicrob.Chemother.*, **41 Suppl C**, 13-19.

17. **Borriello, S. P. and Barclay, F. E.** 1985. Protection of Hamsters Against *Clostridium difficile* Ileocaecitis by Prior Colonization with Non-Pathogenic Strains. *Journal of Medical Microbiology*, **19**, 339-350.

18. **Borriello, S. P., Davies, H. A., and Barclay, F. E.** 1988a. Detection of Fimbriae Amongst Strains of *Clostridium difficile*. *Fems Microbiology Letters*, **49**, 65-67.

19. **Borriello, S. P., Ketley, J. M., Mitchell, T. J., Barclay, F. E., Welch, A. R., Price, A. B., and Stephen, J.** 1987. *Clostridium difficile* - A Spectrum of Virulence and Analysis of Putative Virulence Determinants in the Hamster Model of Antibiotic-Associated Colitis. *Journal of Medical Microbiology*, **24**, 53-64.

20. **Borriello, S. P., Welch, A. R., Barclay, F. E., and Davies, H. A.** 1988b. Mucosal Association by *Clostridium difficile* in the Hamster Gastrointestinal-Tract. *Journal of Medical Microbiology*, **25**, 191-196.

21. **Borriello, S. P., Wren, B. W., Hyde, S., Seddon, S. V., Sibbons, P., Krishna, M. M., Tabaqchali, S., Manek, S., and Price, A. B.** 1992. Molecular, immunological, and biological characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile*. *Infect.Immun.*, **60**, 4192-4199.

22. **Boschetti, E. and Righetti, P. G.** 2008. The ProteoMiner in the proteomic arena: A non-depleting tool for discovering low-abundance species. *Journal of Proteomics*, **71**, 255-264.

23. **Calabi, E. and Fairweather, N.** 2002. Patterns of sequence conservation in the S-layer proteins and related sequences in *Clostridium difficile*. *Journal of Bacteriology*, **184**, 3886-3897.

24. **Calabi, E., Ward, S., Wren, B., Paxton, T., Panico, M., Morris, H., Dell, A., Dougan, G., and Fairweather, N.** 2001. Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Molecular Microbiology*, **40**, 1187-1199.

25. **Castagna, A., Cecconi, D., Sennels, L., Rappsilber, J., Guerrier, L., Fortis, F., Boschetti, E., Lomas, L., and Rigetti, P. G.** 2005. Exploring the hidden human urinary proteome via ligand library beads. *Journal of Proteome Research*, **4**, 1917-1930.

26. **Cordwell, S. J., Nouwens, A. S., and Walsh, B. J.** 2001. Comparative proteomics of bacterial pathogens. *Proteomics*, **1**, 461-472.

27. **Darling, A. E., Mau, B., and Perna, N. T.** 2010. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *Plos One*, **5**,

28. **Davies, H. A. and Borriello, S. P.** 1990. Detection of Capsule in Strains of *Clostridium difficile* of Varying Virulence and Toxigenicity. *Microbial Pathogenesis*, **9**, 141-146.

29. **Deneve, C., Janoir, C., Poilane, I., Fantinato, C., and Collignon, A.** 2009. New trends in *Clostridium difficile* virulence and pathogenesis. *International Journal of Antimicrobial Agents*, **33**, S24-S28.

30. **Driks, A.** 1999. *Bacillus subtilis* spore coat. *Microbiology and Molecular Biology Reviews*, **63**, 1-+.

31. **Drudy, D., Calabi, E., Kyne, L. N. F., Sougioultzis, S., Kelly, E., Fairweather, N., and Kelly, C. P.** 2004. Human antibody response to surface layer proteins in *Clostridium difficile* infection. *Fems Immunology and Medical Microbiology*, **41**, 237-242.

32. **Dupuy, B., Govind, R., Antunes, A., and Matamouros, S.** 2008. *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. *Journal of Medical Microbiology*, **57**, 685-689.

33. **Dupuy, B. and Sonenshein, A. L.** 1998. Regulated transcription of *Clostridium difficile* toxin genes. *Mol.Microbiol.*, **27**, 107-120.

34. **Dworkin, J. and Blaser, M. J.** 1997. Molecular mechanisms of *Campylobacter fetus* surface layer protein expression. *Molecular Microbiology*, **26**, 433-440.

35. **Echan, L. A., Tang, H. Y., Ali-Khan, N., Lee, K., and Speicher, D. W.** 2005. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics*, **5**, 3292-3303.

36. **Encheva, V., Shah, H. N., and Gharbia, S. E.** 2009. Proteomic analysis of the adaptive response of *Salmonella enterica* serovar Typhimurium to growth under anaerobic conditions. *Microbiology-Sgm*, **155**, 2429-2441.

37. **Encheva, V., Wait, R., Gharbia, S. E., Begum, S., and Shah, H. N.** 2005. Proteome analysis of serovars Typhimurium and Pullorum of *Salmonella enterica* subspecies I. *BMC.Microbiol.*, **5**, 42-

38. **Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam, L. T., Buttner, K., Buurman, G., Scharf, C., Venz, S., Volker, U., and Hecker, M.** 2004. A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics*, **4**, 2849-2876.

39. **Fawley, W. N., Underwood, S., Freeman, J., Baines, S. D., Saxton, K., Stephenson, K., Owens, R. C., and Wilcox, M. H.** 2007. Efficacy of hospital cleaning agents and germicides against epidemic *Clostridium difficile* strains. *Infection Control and Hospital Epidemiology*, **28**, 920-925.

40. **Fortis, F., Guerrier, L., Areces, L. B., Antonioli, P., Hayes, T., Carrick, K., Hammond, D., Boschetti, E., and Righetti, P. G.** 2006. A new approach for the detection and identification of protein impurities using combinatorial solid phase ligand libraries. *Journal of Proteome Research*, **5**, 2577-2585.

41. **Fountoulakis, M. and Takacs, B.** 1998. Design of protein purification pathways: Application to the proteome of *Haemophilus influenzae* using heparin chromatography. *Protein Expression and Purification*, **14**, 113-119.

42. **Fountoulakis, M., Takacs, M. F., Berndt, P., Langen, H., and Takacs, B.** 1999a. Enrichment of low abundance proteins of *Escherichia coli* by hydroxyapatite chromatography. *Electrophoresis*, **20**, 2181-2195.

43. **Fountoulakis, M., Takacs, M. F., and Takacs, B.** 1999b. Enrichment of low-copy-number gene products by hydrophobic interaction chromatography. *Journal of Chromatography A*, **833**, 157-168.

44. **Freeman, J., Bauer, M. P., Baines, S. D., Corver, J., Fawley, W. N., Goorhuis, B., Kuijper, E. J., and Wilcox, M. H.** 2010. The Changing Epidemiology of *Clostridium difficile* Infections. *Clinical Microbiology Reviews*, **23**, 529-+.

45. **Geric, B., Johnson, S., Gerding, D. N., Grabnar, M., and Rupnik, M.** 2003. Frequency of binary toxin genes among *Clostridium difficile* strains that do not produce large clostridial toxins. *Journal of Clinical Microbiology*, **41**, 5227-5232.

46. **Gilmore, M. E., Dandyopadhyay, D., Dean, A. M., Linnstaedt, S. D., and Popham, D. L.** 2004. Production of muramic beta-lactam in *Bacillus subtilis* spore peptidoglycan. *Journal of Bacteriology*, **186**, 80-89.

47. **Guerrier, L., Claverol, S., Finzi, L., Paye, F., Fortis, F., Boschetti, E., and Housset, C.** 2007. Contribution of solid-phase hexapeptide ligand libraries to the repertoire of human bile proteins. *Journal of Chromatography A*, **1176**, 192-205.

48. **Hall, I. C. and O'Toole, E.** 1935. Intestinal flora in new-borin infants - With a description of a new pathogenic anaerobe, *Bacillus difficilis* 2. *American Journal of Diseases of Children*, **49**, 390-402.

49. **Hammond, G. A. and Johnson, J. L.** 1995. The Toxigenic Element of *Clostridium difficile* Strain Vpi-10463. *Microbial Pathogenesis*, **19**, 203-213.

50. **Haslam, S. C., Ketley, J. M., Mitchell, T. J., Stephen, J., Burdon, D. W., and Candy, D. C. A.** 1986. Growth of *Clostridium difficile* and Production of Toxins A and B in Complex and Defined Media. *Journal of Medical Microbiology*, **21**, 293-297.

51. **He, M., Sebaihia, M., Lawley, T. D., Stabler, R. A., Dawson, L. F., Martin, M. J., Holt, K. E., Seth-Smith, H. M. B., Quail, M. A., Rance, R., Brooks, K., Churcher, C., Harris, D., Bentley, S. D., Burrows, C., Clark, L., Corton, C., Murray, V., Rose, G., Thurston, S., van Tonder, A., Walker, D., Wren, B. W., Dougan, G., and Parkhill, J.** 2010. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 7527-7532.

52. **Hecht, G., Koutsouris, A., Pothoulakis, C., Lamont, J. T., and Madara, J. L.** 1992. *Clostridium difficile* Toxin-B Disrupts the Barrier Function of T84 Monolayers. *Gastroenterology*, **102**, 416-423.

53. **Hecker, M., Antelmann, H., Buettner, K., and Bernhardt, J.** 2008. Gel-based proteomics of Gram-positive bacteria: A powerful tool to address physiological questions. *Proteomics*, **8**, 4958-4975.

54. **Henderson, B., Allan, E., and Coates, A. R. M.** 2006. Stress wars: the direct role of host and bacterial molecular chaperones in bacterial infection. *Infection and Immunity*, **74**, 3693-3706.

55. **Hennequin, C., Porcheray, F., Waligora-Dupriet, A. J., Collignon, A., Bare, M. C., Bourlioux, P., and Karjalainen, T.** 2001. GroEL (Hsp60) of *Clostridium difficile* is involved in cell adherence. *Microbiology-Uk*, **147**, 87-96.

56. **Hirabayashi, J.** 2008. Concept, strategy and realization of lectin-based glycan profiling. *Journal of Biochemistry*, **144**, 139-147.

57. **Hubert, B., Loo, V. G., Bourgault, A. M., Poirier, L., Dascal, A., Fortin, E., Dionne, M., and Lorange, M.** 2007. A portrait of the geographic dissemination of the *Clostridium difficile* North American pulsed-field type 1 strain and the epidemiology of C-difficile-associated disease in Quebec. *Clinical Infectious Diseases*, **44**, 238-244.

58. **Hundsberger, T., Braun, V., Weidmann, M., Leukel, P., Sauerborn, M., and vonEichelStreiber, C.** 1997. Transcription analysis of the genes tcdA-E of the pathogenicity locus of *Clostridium difficile*. *European Journal of Biochemistry*, **244**, 735-742.

59. **James, P.** 1997. Protein identification in the post-genome era: the rapid rise of proteomics. *Quarterly Reviews of Biophysics*, **30**, 279-331.

60. **Karjalainen, T., Waligora-Dupriet, A. J., Cerquetti, M., Spigaglia, P., Maggioni, A., Mauri, P., and Mastrantonio, P.** 2001. Molecular and genomic analysis of genes encoding surface-anchored proteins from *Clostridium difficile*. *Infection and Immunity*, **69**, 3442-3446.

61. **Karlsson, S., Dupuy, B., Mukherjee, K., Norin, E., Burman, L. G., and Akerlund, T.** 2003. Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. *Infection and Immunity*, **71**, 1784-1793.

62. **Karp, N. A., Feret, R., Rubtsov, D. V., and Lilley, K. S.** 2008. Comparison of DIGE and post-stained gel electrophoresis with both traditional and SameSpots analysis for quantitative proteomics. *Proteomics*, **8**, 948-960.

63. **Keidel, E. M., Ribitsch, D., and Lottspeich, F.** 2010. Equalizer technology - Equal rights for disparate beads. *Proteomics*, **10**, 2089-2098.

64. **Kelly, C. P. and Lamont, J. T.** 1998. *Clostridium difficile* infection. *Annual Review of Medicine*, **49**, 375-390.

65. **Kirby, J. M., Ahern, H., Roberts, A. K., Kumar, V., Freeman, Z., Acharya, K. R., and Shone, C. C.** 2009. Cwp84, a Surface-associated Cysteine Protease, Plays a Role in the Maturation of the Surface Layer of *Clostridium difficile*. *Journal of Biological Chemistry*, **284**, 34666-34673.

66. **Krishna, M. M., Powell, N. B. L., and Borriello, S. P.** 1996. Cell surface properties of *Clostridium difficile*: Haemagglutination, relative hydrophobicity and charge. *Journal of Medical Microbiology*, **44**, 115-123.

67. **Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A., and Minton, N. P.** 2010. The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature*, **467**, 711-U97.

68. **Kyne, L., Warny, M., Qamar, A., and Kelly, C. P.** 2001. Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea. *Lancet*, **357**, 189-193.

69. **Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., and Knapp, R. J.** 1991. A New Type of Synthetic Peptide Library for Identifying Ligand-Binding Activity. *Nature*, **354**, 82-84.

70. **Larson, H. E., Barclay, F. E., Honour, P., and Hill, I. D.** 1982. Epidemiology of *Clostridium difficile* in Infants. *Journal of Infectious Diseases*, **146**, 727-733.

71. **Larson, H. E. and Borriello, S. P.** 1990. Quantitative Study of Antibiotic-Induced Susceptibility to *Clostridium difficile* Enterocecitis in Hamsters. *Antimicrobial Agents and Chemotherapy*, **34**, 1348-1353.

72. **Larson, H. E., Price, A. B., Honour, P., and Borriello, S. P.** 1978. *Clostridium difficile* and Etiology of Pseudomembranous Colitis. *Lancet*, **1**, 1063-1066.

73. **Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J., and Dougan, G.** 2009. Proteomic and Genomic Characterization of Highly Infectious *Clostridium difficile* 630 Spores. *Journal of Bacteriology*, **191**, 5377-5386.

74. **Leon-Kempis, M. D., Guccione, E., Mulholland, F., Williamson, M. P., and Kelly, D. J.** 2006. The *Campylobacter jejuni* PEB1a adhesin is an aspartate/glutamate-binding protein of an ABC transporter essential for

microaerobic growth on dicarboxylic amino acids. *Molecular Microbiology*, **60**, 1262-1275.

75. **Li, Ming Shi, Chow, Noel Y., Sinha, Sunita, Halliwell, Denise, Finney, Michelle, Gorringe, Andrew R., Watson, Mark W., Kroll, J., Langford, Paul R., and Webb, Steven A.** 2009. A *Neisseria meningitidis* NMB1966 mutant is impaired for invasion of respiratory epithelial cells, survival in human blood and for virulence in vivo. *Medical Microbiology and Immunology*, **198**, 57-67.

76. **Li, S. M., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D. J., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q. R., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H. Y., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., van den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E., and Vidal, M.** 2004. A map of the interactome network of the metazoan *C-elegans*. *Science*, **303**, 540-543.

77. **Lima, A. A. M., Lyerly, D. M., Wilkins, T. D., Innes, D. J., and Guerrant, R. L.** 1988. Effects of *Clostridium difficile* Toxin-A and Toxin-B in Rabbit Small and Large-Intestine *In vivo* and on Cultured-Cells *In vitro*. *Infection and Immunity*, **56**, 582-588.

78. **Logan, S. M.** 2006. Flagellar glycosylation - a new component of the motility repertoire? *Microbiology-Sgm*, **152**, 1249-1262.

79. **Loo, V. G.** 2006. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality (vol 353, pg 2442, 2005). *New England Journal of Medicine*, **354**, 2200-2200.

80. **Lyerly, D. M., Saum, K. E., Macdonald, D. K., and Wilkins, T. D.** 1985. Effects of *Clostridium difficile* Toxins Given Intragastrically to Animals. *Infection and Immunity*, **47**, 349-352.

81. **Lyras, D., O'Connor, J. R., Howarth, P. M., Sambol, S. P., Carter, G. P., Phumoonna, T., Poon, R., Adams, V., Vedantam, G., Johnson, S.,**

Gerding, D. N., and Rood, J. I. 2009. Toxin B is essential for virulence of *Clostridium difficile*. *Nature*, **458**, 1176-1181.

82. **McCoubrey, J. and Poxton, I. R.** 2001. Variation in the surface layer proteins of *Clostridium difficile*. *Fems Immunology and Medical Microbiology*, **31**, 131-135.

83. **McDonald, L. C., Killgore, G. E., Thompson, A., Owens, R. C., Kazakova, S. V., Sambol, S. P., Johnson, S., and Gerding, D. N.** 2005. An epidemic, toxin gene-variant strain of *Clostridium difficile*. *New England Journal of Medicine*, **353**, 2433-2441.

84. **Merrifield, R. B.** 1965. Automated Synthesis of Peptides. *Science*, **150**, 178-&.

85. **Merrigan, M., Venugopal, A., Mallozzi, M., Roxas, B., Viswanathan, V. K., Johnson, S., Gerding, D. N., and Vedantam, G.** 2010. Human Hypervirulent *Clostridium difficile* Strains Exhibit Increased Sporulation as Well as Robust Toxin Production. *Journal of Bacteriology*, **192**, 4904-4911.

86. **Mitchell, T. J., Ketley, J. M., Haslam, S. C., Stephen, J., Burdon, D. W., Candy, D. C. A., and Daniel, R.** 1986. Effect of Toxin-A and Toxin-B of *Clostridium difficile* on Rabbit Ileum and Colon. *Gut*, **27**, 78-85.

87. **Moncrief, J. S., Barroso, L. A., and Wilkins, T. D.** 1997. Positive regulation of *Clostridium difficile* toxins. *Infection and Immunity*, **65**, 1105-1108.

88. **Mukherjee, K., Karlsson, S., Burman, L. G., and Akerlund, T.** 2002. Proteins released during high toxin production in *Clostridium difficile*. *Microbiology*, **148**, 2245-2253.

89. **Mulligan, M. E., Miller, S. D., Mcfarland, L. V., Fung, H. C., and Kwok, R. Y. Y.** 1993. Elevated Levels of Serum Immunoglobulins in Asymptomatic Carriers of *Clostridium difficile*. *Clinical Infectious Diseases*, **16**, S239-S244.

90. **O'Connor, J. R., Johnson, S., and Gerding, D. N.** 2009. *Clostridium difficile* Infection Caused by the Epidemic BI/NAP1/027 Strain. *Gastroenterology*, **136**, 1913-1924.

91. **Pechine, S., Gleizes, A., Janoir, C., Gorges-Kergot, R., Barc, M. C., Delmee, M., and Collignon, A.** 2005a. Immunological properties of surface proteins of *Clostridium difficile*. *Journal of Medical Microbiology*, **54**, 193-196.

92. **Pechine, S., Janoir, C., and Collignon, A.** 2005b. Variability of *Clostridium difficile* surface proteins and specific serum antibody response in patients with *Clostridium difficile*-associated disease. *Journal of Clinical Microbiology*, **43**, 5018-5025.

93. **Pernemalm, M., Lewensohn, R., and Lehtio, J.** 2009. Affinity prefractionation for MS-based plasma proteomics. *Proteomics*, **9**, 1420-1427.

94. **Pieper, R., Su, Q., Gatlin, C. L., Huang, S. T., Anderson, N. L., and Steiner, S.** 2003. Multi-component immunoaffinity subtraction chromatography: An innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics*, **3**, 422-432.

95. **Piggot, P. J. and Hilbert, D. W.** 2004. Sporulation of *Bacillus subtilis*. *Current Opinion in Microbiology*, **7**, 579-586.

96. **Popoff, M. R., Rubin, E. J., Gill, D. M., and Boquet, P.** 1988. Actin-Specific Adp-Ribosyltransferase Produced by A *Clostridium difficile* Strain. *Infection and Immunity*, **56**, 2299-2306.

97. **Ramos, H. C., Rumbo, M., and Sirard, J. C.** 2004. Bacterial flagellins: mediators of pathogenicity and host immune responses in mucosa. *Trends in Microbiology*, **12**, 509-517.

98. **Rathsam, C., Eaton, R. E., Simpson, C. L., Browne, G. V., Valova, V. A., Harty, D. W. S., and Jacques, N. A.** 2005. Two-dimensional fluorescence difference gel electrophoretic analysis of *Streptococcus mutans* biofilms. *Journal of Proteome Research*, **4**, 2161-2173.

99. **Riedel, K., Carranza, P., Gehrig, P., Potthast, F., and Eberl, L.** 2006. Towards the proteome of *Burkholderia cenocepacia* H111: Setting up a 2-DE reference map. *Proteomics*, **6**, 207-216.

100. **Righetti, P. G. and Boschetti, E.** 2008. The Proteominer and the Fortyniners: Searching for Gold Nuggets in the Proteomic Arena. *Mass Spectrometry Reviews*, **27**, 596-608.

101. **Righetti, P. G., Castagna, A., Herbert, B., Reymond, F., and Rossier, J. S.** 2003. Prefractionation techniques in proteome analysis. *Proteomics*, **3**, 1397-1407.

102. **Rupnik, M.** 2008. Heterogeneity of large clostridial toxins: importance of *Clostridium difficile* toxinotypes. *Fems Microbiology Reviews*, **32**, 541-555.

103. **Rupnik, M., Wilcox, M. H., and Gerding, D. N.** 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nature Reviews Microbiology*, **7**, 526-536.

104. **Sanchez-Hurtado, K., Corretge, M., Mutlu, E., McIlhagger, R., Starr, J. M., and Poxton, I. R.** 2008. Systemic antibody response to *Clostridium difficile* in colonized patients with and without symptoms and matched controls. *Journal of Medical Microbiology*, **57**, 717-724.

105. **Sara, M. and Sleytr, U. B.** 2000. S-layer proteins. *Journal of Bacteriology*, **182**, 859-868.

106. **Savariau-Lacomme, M. P., Lebarbier, C., Karjalainen, T., Collignon, A., and Janoir, C.** 2003. Transcription and analysis of polymorphism in a cluster of genes encoding surface-associated proteins of *Clostridium difficile*. *Journal of Bacteriology*, **185**, 4461-4470.

107. **Savidge, T. C., Pan, W. H., Newman, P., O'Brien, M., Anton, P. M., and Pothoulakis, C.** 2003. *Clostridium difficile* toxin B is an inflammatory enterotoxin in human intestine. *Gastroenterology*, **125**, 413-420.

108. **Schwan, C., Stecher, B., Tzivelekidis, T., van Ham, M., Rohde, M., Hardt, W. D., Wehland, J., and Aktories, K.** 2009. *Clostridium difficile* Toxin CDT Induces Formation of Microtubule-Based Protrusions and Increases Adherence of Bacteria. *Plos Pathogens*, **5**,

109. **Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdeno-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P.,**

Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabbinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B., and Parkhill, J. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat.Genet.*, **38**, 779-786.

110. **Seddon, S. V., Hemingway, I., and Borriello, S. P.** 1990. Hydrolytic Enzyme-Production by *Clostridium difficile* and Its Relationship to Toxin Production and Virulence in the Hamster Model. *Journal of Medical Microbiology*, **31**, 169-174.

111. **Sennels, L., Salek, M., Lomas, L., Boschetti, E., Righetti, P. G., and Rappsilber, J.** 2007. Proteomic analysis of human blood serum using peptide library beads. *Journal of Proteome Research*, **6**, 4055-4062.

112. **Serrano, M., Corte, L., Opdyke, J., Moran, C. P., and Henriques, A. O.** 2003. Expression of spoIIIJ in the prespore is sufficient for activation of sigma(G) and for sporulation in *Bacillus subtilis*. *Journal of Bacteriology*, **185**, 3905-3917.

113. **Setlow, P.** 2007. I will survive: DNA protection in bacterial spores. *Trends in Microbiology*, **15**, 172-180.

114. **Shah, H. N.** 2010. Microbial Characterisation; the Transition from Conventional Methods to Proteomics. In *Mass Spectrometry for Microbial Proteomics* (Shah H. N., Gharbia S. E. Eds), p.339. John Wiley and Sons Ltd, West Sussex.

115. **Sharp, J. and Poxton, I. R.** 1986. Analysis of the Membrane Lipocarbohydrate Antigen of *Clostridium difficile* by Polyacrylamide-Gel Electrophoresis and Immunoblotting. *Fems Microbiology Letters*, **34**, 97-100.

116. **Sihlbom, C., Kanmert, I., von Bahr, H., and Davidsson, P.** 2008. Evaluation of the combination of bead technology with SELDI-TOF-MS and 2-D DIGE for detection of plasma proteins. *Journal of Proteome Research*, **7**, 4191-4198.

117. **Smith A.** 30-6-2005. Outbreak of *Clostridium difficile* infection in in an English hospital linked to hypertoxin -producing strains in Canada and the US. *Eurosurveillance*, **10**,

118. **Sougioultzis, S., Kyne, L., Drudy, D., Keates, S., Maroo, S., Pothoulakis, C., Giannasca, P. J., Lee, C. K., Warny, M., Monath, T. P., and Kelly, C. P.** 2005. *Clostridium difficile* toxoid vaccine in recurrent *C-difficile*-associated diarrhea. *Gastroenterology*, **128**, 764-770.

119. **Spigaglia, P. and Mastrantonio, P.** 2004. Comparative analysis of *Clostridium difficile* clinical isolates belonging to different genetic lineages and time periods. *Journal of Medical Microbiology*, **53**, 1129-1136.

120. **Stabler, R. A., Gerding, D. N., Songer, J. G., Drudy, D., Brazier, J. S., Trinh, H. T., Witney, A. A., Hinds, J., and Wren, B. W.** 2006. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. *J.Bacteriol.*, **188**, 7297-7305.

121. **Stabler, R. A., He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., Lawley, T. D., Sebaihia, M., Quail, M. A., Rose, G., Gerding, D. N., Gibert, M., Popoff, M. R., Parkhill, J., Dougan, G., and Wren, B. W.** 2009. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biology*, **10**,

122. **Sullivan, L. and Bennett, G. N.** 2006. Proteome analysis and comparison of *Clostridium acetobutylicum* ATCC 824 and Spo0A strain variants. *Journal of Industrial Microbiology & Biotechnology*, **33**, 298-308.

123. **Sullivan, N. M., Pellett, S., and Wilkins, T. D.** 1982. Purification and Characterization of Toxin-A and Toxin-B of *Clostridium difficile*. *Infection and Immunity*, **35**, 1032-1040.

124. **Sun, X. S., Chiu, J. F., and He, Q. Y.** 2005. Application of immobilized metal affinity chromatography in proteomics. *Expert Review of Proteomics*, **2**, 649-657.

125. **Tan, K. S., Wee, B. Y., and Song, K. P.** 2001. Evidence for holin function of tcdE gene in the pathogenicity of *Clostridium difficile*. *Journal of Medical Microbiology*, **50**, 613-619.

126. **Tasteyre, A., Barc, M. C., Collignon, A., Boureau, H., and Karjalainen, T.** 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. *Infection and Immunity*, **69**, 7937-7940.

127. **Tasteyre, A., Barc, M. C., Karjalainen, T., Dodson, P., Hyde, S., Bourlioux, P., and Borriello, P.** 2000a. A *Clostridium difficile* gene encoding flagellin. *Microbiology-Sgm*, **146**, 957-966.

128. **Tasteyre, A., Karjalainen, T., Avesani, V., Delmee, M., Collignon, A., Bourlioux, P., and Barc, M. C.** 2000b. Phenotypic and genotypic diversity of the flagellin gene (fliC) among *Clostridium difficile* isolates from different serogroups. *Journal of Clinical Microbiology*, **38**, 3179-3186.

129. **Thulasiraman, V., Lin, S. H., Gheorghiu, L., Lathrop, J., Lomas, L., Hammond, D., and Boschetti, E.** 2005. Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. *Electrophoresis*, **26**, 3561-3571.

130. **Twine, S. M., Reid, C. W., Aubry, A., McMullin, D. R., Fulton, K. M., Austin, J., and Logan, S. M.** 2009. Motility and Flagellar Glycosylation in *Clostridium difficile*. *Journal of Bacteriology*, **191**, 7050-7062.

131. **Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., Wilcox, M. H., and Stephenson, K.** 2009. Characterization of the Sporulation Initiation Pathway of *Clostridium difficile* and Its Role in Toxin Production. *Journal of Bacteriology*, **191**, 7296-7305.

132. **Unlu, M., Morgan, M. E., and Minden, J. S.** 1997. Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis*, **18**, 2071-2077.

133. **van Alphen, L. B., Wuhrer, M., Bleumink-Pluym, N. M. C., Hensbergen, P. J., Deelder, A. M., and van Putten, J. P. M.** 2008. A functional *Campylobacter jejuni* maf4 gene results in novel glycoforms on flagellin and altered autoagglutination behaviour. *Microbiology-Sgm*, **154**, 3385-3397.

134. **Van de Velde, S., Delaive, E., Dieu, M., Carryn, S., Van Bambeke, F., Devreese, B., Raes, M., and Tulkens, P. M.** 2009. Isolation and 2-D-DIGE proteomic analysis of intracellular and extracellular forms of *Listeria monocytogenes*. *Proteomics*, **9**, 5484-5496.

135. **Voth, D. E. and Ballard, J. D.** 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. *Clin.Microbiol.Rev.*, **18**, 247-263.

136. **Waligora, A. J., Barc, M. C., Bourlioux, P., Collignon, A., and Karjalainen, T.** 1999. *Clostridium difficile* cell attachment is modified by environmental factors. *Applied and Environmental Microbiology*, **65**, 4234-4238.

137. **Waligora, A. J., Hennequin, C., Mullany, P., Bourlioux, P., Collignon, A., and Karjalainen, T.** 2001. Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. *Infection and Immunity*, **69**, 2144-2153.

138. **Wang, Q. F., Suzuki, A., Mariconda, S., Porwollik, S., and Harshey, R. M.** 2005. Sensing wetness: a new role for the bacterial flagellum. *Embo Journal*, **24**, 2034-2042.

139. **Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J., Frost, E., and McDonald, L. C.** 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *Lancet*, **366**, 1079-1084.

140. **Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., Yan, J. X., Gooley, A. A., Hughes, G., HumpherySmith, I., Williams, K. L., and Hochstrasser, D. F.** 1996. From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Bio-Technology*, **14**, 61-65.

141. **Wright, A., Drudy, D., Kyne, L., Brown, K., and Fairweather, N. F.** 2008. Immunoreactive cell wall proteins of *Clostridium difficile* identified by human sera. *Journal of Medical Microbiology*, **57**, 750-756.

142. **Wright, A., Wait, R., Begum, S., Crossett, B., Nagy, J., Brown, K., and Fairweather, N.** 2005. Proteomic analysis of cell surface proteins from *Clostridium difficile*. *Proteomics*, **5**, 2443-2452.

143. **Yamakawa, K., Karasawa, T., Ohta, T., Hayashi, H., and Nakamura, S.** 1998. Inhibition of enhanced toxin production by *Clostridium difficile* in biotin-limited conditions. *Journal of Medical Microbiology*, **47**, 767-771.

Appendices

Table A1 – *Identifications of protein spots picked from 2D reference maps with MASCOT identification scores (A), theoretical mass (B) and pI (C), percentage peptide coverage (D) and number of matched peptides (E). Any spot number not shown did not give rise to an identification.*

Strain A							
Spot	**Protein**	**ID no**	**A**	**B**	**C**	**D**	**E**
A1	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269	sta_BASYS03269	141	80379	4.83	40	28
A2	atpA, 68761-67259 (CounterClockwise) ATP synthase subunit alpha	sta_BASYS00075	70	54688	4.91	24	11
A3	stage IV sporulation protein A	630_gi|115251680|emb|CAJ69515.1|	77	56373	4.97	32	13
A5	electron transfer flavoprotein alpha-subunit	630_gi|115249407|emb|CAJ67222.1|	125	57353	5.03	51	15
A6	(R)-2-hydroxyisocaproate dehydrogenase	630_gi|115249400|emb|CAJ67215.1|	108	36552	5.07	55	14
A7	electron transfer flavoprotein beta-subunit	630_gi|115249406|emb|CAJ67221.1|	131	28767	4.87	78	15
A8	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	105	64664	4.76	41	16
A10	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269	sta_BASYS03269	88	80379	4.83	28	15
A11	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405|emb|CAJ67220.1|	89	41009	6.41	40	10
A12	caiA, 763783-764928 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	sta_BASYS00745	170	41469	5.88	69	25
B1	ydbK, 3330872-3327333 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase)	sta_BASYS03153	80	130112	5.39	21	20
B2	ydbK, 3330872-3327333 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase)	sta_BASYS03153	116	130112	5.39	28	29
B3	ydbK, 3330872-3327333 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase)	630_gi|115251733|emb|CAJ69568.1|	137	130112	5.39	29	30
B4	ydbK, 3330872-3327333 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase)	sta_BASYS03153	108	130112	5.39	25	24
B5	ydbK, 3330872-3327333 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase)	sta_BASYS03153	111	130112	5.39	26	26

B6	tyrB, 2974977-2973718 (CounterClockwise) Hypothetical Protein tyrB (putative aspartate aminotransferase)	sta_BASYS02834	77	47850	5.17	28	13
B9	isoleucyl-tRNA synthetase	630_gi|115251669|emb|CAJ69504.1|	135	120425	5.08	21	18
B10	rpoB, 408612-412328 (Clockwise) Hypothetical Protein rpoB (DNA-directed RNA polymerase beta chain)	sta_BASYS00405	86	139806	4.76	20	20
C5	infB, 1691221-1693161 (Clockwise) Translation initiation factor IF-2 (translation initiation factor IF-2)	sta_BASYS01609	84	69895	4.76	30	12
C6	dnaK, 3071723-3069876 (CounterClockwise) Hypothetical Protein dnaK (chaperone protein)	sta_BASYS02928	102	66541	4.72	33	15
C7	groL, 567992-569620 (Clockwise) Hypothetical Protein groL (60 kDa chaperonin)	sta_BASYS00550	109	37735	4.74	50	14
C10	fusA, 417208-419274 (Clockwise) Hypothetical Protein fusA	sta_BASYS00409	54	76076	4.91	24	12
D2	(tuf1, 402179-403372 (Clockwise) Elongation factor Tu (elongation factor TU)	sta_BASYS00396	159	44226	4.94	49	19
D3	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269	sta_BASYS03269	128	80379	4.83	33	19
D10	aegA, 1953842-1955236 (Clockwise) Hypothetical Protein aegA (putative glutamate synthase NADPH small chain)	sta_BASYS01860	75	50248	5.17	31	11
D11	atpB, 3616679-3615306 (CounterClockwise) V-type ATP synthase beta chain (V-type sodium ATP synthase subunit B)	sta_BASYS03388	79	50585	5.21	43	15
D12	atpD, 66340-64946 (CounterClockwise) ATP synthase subunit beta (ATP synthase beta chain)	sta_BASYS00073	141	49817	5.18	62	19
E1	yjiR, 302052-300856 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	sta_BASYS00306	61	44997	5.12	30	10
E2	yjiR, 302052-300856 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	sta_BASYS00306	160	44997	5.12	65	25
E3	NAD-specific glutamate dehydrogenase	630_gi|115249189|emb|CAJ67001.1|	81	46271	5.16	42	15
	putative amino acid aminotransferase	630_gi|115252729|emb|CAJ70573.1|	63	44997	5.12	37	14
E4	frc, 759341-760540 (Clockwise) Hypothetical Protein frc (isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase)	sta_BASYS00741	87	44491	5.18	43	12
E5	atoB, 1445216-1446448 (Clockwise) Hypothetical Protein atoB (acetyl-CoA acetyltransferase)	sta_BASYS01347	78	41005	5.56	40	13
E8	yceH, 2127930-2129066 (Clockwise) Uncharacterized protein yceH (putative tellurite resistance protein)	sta_BASYS02042	106	43643	5.09	39	13
E10	BASYS00469, 475249-476328 (Clockwise) Hypothetical Protein BASYS00469 (putative oxidoreductase, thiamine diP-binding subunit)	sta_BASYS00469	92	39290	5.32	54	12
E12	ackA, 473216-474295 (Clockwise) Hypothetical Protein ackA (butyrate kinase)	sta_BASYS00466	130	38728	5.34	67	23

F1	yjiM, 762620-763747 (Clockwise) Hypothetical Protein yjiM (subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase)	sta_BASYS00744	62	42993	5.27	35	12
F5	ABC transporter, substrate-binding lipoprotein	630_gi|115249887|emb|CAJ67706.1|	62	36159	5.4	36	7
F6	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269	sta_BASYS03269	86	80379	4.83	26	18
F7	agaY, 769977-769051 (CounterClockwise) Hypothetical Protein agaY (putative fructose-bisphosphate aldolase)	sta_BASYS00749	67	33069	5.16	46	9
F8	ackA, 2971240-2970170 (CounterClockwise) Hypothetical Protein ackA (butyrate kinase)	sta_BASYS02831	73	38869	5.29	34	8
F9	cysM, 2076886-2077794 (Clockwise) Hypothetical Protein cysM (putative O-acetylserine sulfhydrylase)	sta_BASYS01991	102	32816	5.35	57	13
F10	cysM, 2076886-2077794 (Clockwise) Hypothetical Protein cysM (putative O-acetylserine sulfhydrylase)	sta_BASYS01991	98	32816	5.35	58	12
F11	prs, 113742-112792 (CounterClockwise) Hypothetical Protein prs (ribose-phosphate pyrophosphokinase)	sta_BASYS00121	72	34548	5.41	36	8
G7	BASYS01569, 1654445-1655230 (Clockwise) Hypothetical Protein BASYS01569 (GTP-sensing transcriptional pleiotropic repressor)	sta_BASYS01569	92	28727	5.31	48	11
H3	BASYS02832, 2971868-2971293 (CounterClockwise) Hypothetical Protein BASYS02832 (putative indolepyruvate oxidoreductase subunit)	sta_BASYS02832	85	21020	5.44	68	9
H7	rplJ, 406008-406538 (Clockwise) Hypothetical Protein rplJ (50S ribosomal protein L10)	sta_BASYS00402	72	18582	6.2	56	9
H8	BASYS03154, 3331959-3331390 (CounterClockwise) Propanediol Utilization Protein	sta_BASYS03154	63	20079	5.97	52	8
I1	adhE, 3632084-3629430 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase)	sta_BASYS03400	206	97197	5.68	47	34
I7	pflB, 1174270-1172039 (CounterClockwise) Hypothetical Protein pflB (formate acetyltransferase)	sta_BASYS01106	169	83966	5.39	46	29
I8	formate acetyltransferase	630_gi|115249776|emb|CAJ67593.1|	109	83966	5.39	46	27
I8	polyribonucleotide nucleotidyltransferase	630_gi|115250354|emb|CAJ68176.1|	99	78172	5.3	40	23
J1	acpS, 64221-63841 (CounterClockwise) Hypothetical Protein acpS (holo-acyl-carrier protein synthase)	sta_BASYS00071	56	14473	8.63	48	6
J3	BASYS01060, 1118805-1120724 (Clockwise) Hypothetical Protein BASYS01060 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	sta_BASYS01060	69	69693	5.5	25	12

J4	BASYS01060, 1118805-1120724 (Clockwise) Hypothetical Protein BASYS01060 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	sta_BASYS01060	208	69709	5.5	53	31
J5	fusA, 353018-354958 (Clockwise) Hypothetical Protein fusA (putative translation elongation factor)	sta_BASYS00349	175	72577	5.45	58	31
J6	fusA, 353018-354958 (Clockwise) Hypothetical Protein fusA (putative translation elongation factor)	sta_BASYS00349	119	72577	5.45	43	21
J7	arginyl-tRNA synthetase	630_gi|115249727|emb|CAJ67544.1|	88	65234	5.38	36	14
J7	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184|emb|CAJ66996.1|	67	69298	5.41	27	13
J10	indolepyruvate oxidoreductase subunit	630_gi|115251433|emb|CAJ69266.1|	195	68929	5.65	49	26
J12	BASYS03267, 3473846-3471975 (CounterClockwise) Hypothetical Protein BASYS03267 (cell surface protein (putative S-layer protein precursor))	sta_BASYS03267	155	66427	6.2	47	22
K1	BASYS01063, 1122044-1123738 (Clockwise) Hypothetical Protein BASYS01063 (formate--tetrahydrofolate ligase)	sta_BASYS01063	162	60894	5.53	56	26
K2	thiH, 2730080-2731456 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	sta_BASYS02617	94	52653	5.56	43	17
K3	thiH, 2730080-2731456 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	sta_BASYS02617	103	52653	5.56	51	19
K4	thiH, 2730080-2731456 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	sta_BASYS02617	84	52653	5.56	38	12
K5	glucose-6-phosphate isomerase	630_gi|115252341|emb|CAJ70182.1|	86	50677	5.48	50	15
K6	putative transcriptional regulator	630_gi|115252733|emb|CAJ70577.1|	75	50203	5.56	39	18
	tyrS2, 1934471-1935679 (Clockwise) Tyrosyl-tRNA synthetase 2 (tyrosyl-tRNA synthetase)	sta_BASYS01844	98	45703	5.42	42	14
K7	thiH, 2730080-2731456 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	sta_BASYS02617	94	52653	5.56	43	17
K9	ackA, 1545979-1547175 (Clockwise) Acetate kinase (acetate kinase)	sta_BASYS01450	100	43619	5.4	56	21
K11	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	630_gi|115249053|emb|CAJ66864.1|	86	59419	5.22	39	16
K11	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	630_gi|115249740|emb|CAJ67557.1|	68	50103	5.27	36	12
L1	BASYS02792, 2934707-2933211 (CounterClockwise) Hypothetical Protein BASYS02792 (gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase)	sta_BASYS02792	74	55864	5.8	36	13

L2	BASYS02792, 2934707-2933211 (CounterClockwise) Hypothetical Protein BASYS02792 (gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase)	sta_BASYS02792	56	55865	5.8	26	9
L5	glyA, 3392427-3391183 (CounterClockwise) Hypothetical Protein glyA (putative serine hydroxymethyltransferase)	sta_BASYS03201	149	46078	5.67	46	16
L7	iscS, 1657961-1659169 (Clockwise) Hypothetical Protein iscS (cysteine desulfurase)	sta_BASYS01574	106	44552	5.9	33	15
L8	caiA, 763783-764928 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	sta_BASYS00745	160	41469	5.88	54	19
L9	caiA, 763783-764928 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	sta_BASYS00745	90	41469	5.88	23	8
L11	ydiO, 1440421-1441557 (Clockwise) Hypothetical Protein ydiO (butyryl-CoA dehydrogenase)	sta_BASYS01342	131	41348	5.7	51	17
L12	gapC, 3929194-3928187 (CounterClockwise) Hypothetical Protein gapC (glyceraldehyde-3-phosphate dehydrogenase 2)	sta_BASYS03662	69	36239	5.75	31	9
M1	paaH, 1444341-1445186 (Clockwise) Hypothetical Protein paaH (3-hydroxybutyryl-CoA dehydrogenase)	sta_BASYS01346	127	30887	5.8	49	13
M2	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	65	34354	6.51	36	8
M3	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	95	34354	6.51	44	12
M4	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	88	34354	6.51	44	13
M5	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	92	34354	6.51	41	12
M6	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	72	34354	6.51	37	8
M7	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	63	34354	6.51	35	8
M8	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	61	34354	6.51	33	6
M9	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	59	34354	6.51	31	6
M11	putative flagellar hook-associated protein	630_gi|115249239|emb|CAJ67052.1|	68	48018	4.94	41	12
M12	metF, 1126186-1127067 (Clockwise) Methylenetetrahydrofolate Reductase (putative methylenetetrahydrofolate reductase)	sta_BASYS01067	63	31776	5.81	30	9
N1	BASYS00470, 476328-477080 (Clockwise) Hypothetical Protein BASYS00470 (putative subunit of oxidoreductase)	sta_BASYS00470	79	27031	6.38	38	8
N3	rpsB, 2712313-2711600 (CounterClockwise) 30S ribosomal protein S2 (30S ribosomal protein S2)	sta_BASYS02601	63	26940	6.19	44	10
N4	yjiL, 760570-761379 (Clockwise) Hypothetical Protein yjiL (activator of 2-hydroxyisocaproyl-CoA dehydratase)	sta_BASYS00742	82	28948	6.35	40	9

N5	spo0A, 1583111-1583935 (Clockwise) Stage 0 sporulation protein A homolog	sta_BASYS01490	119	30888	6.33	57	14
N7	fusA, 417208-419274 (Clockwise) Hypothetical Protein fusA	sta_BASYS01410	46	**48962**	6.3	24	8
N8	clpB, 357405-359852 (Clockwise) Hypothetical Protein clpB (ATP-dependent Clp protease)	sta_BASYS00354	209	91292	5.85	46	32
N9	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	186	64664	4.76	51	24
N10	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03073	stb_BASYS03072	191	64664	4.76	57	27
N12	caiA, 763783-764928 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	sta_BASYS00745	92	41469	5.88	41	13
R1	BASYS01062, 1121034-1121813 (Clockwise) Hypothetical Protein BASYS01062 (putative carbon monoxide dehydrogenase accessory protein)	sta_BASYS01062	98	28728	5.69	50	13
R3	ydiA, 3015428-3014586 (CounterClockwise) Putative phosphotransferase CD2411 (conserved hypothetical protein)	sta_BASYS02873	120	32006	5.95	55	15
R4	BASYS00467, 474264-474962 (Clockwise) Hypothetical Protein CHY (putative ATP/GTP-binding protein)	sta_BASYS00467	57	26529	8.37	33	7
R7	ydiO, 1440421-1441557 (Clockwise) Hypothetical Protein ydiO	sta_BASYS01342	183	41348	5.7	74	22
R8	ypdF, 2942563-2941484 (CounterClockwise) Hypothetical Protein ypdF (putative Xaa-Pro dipeptidase)	sta_BASYS02798	82	40623	5.5	38	11
R11	ygfH, 2932777-2931470 (CounterClockwise) Hypothetical Protein ygfH	sta_BASYS02790	100	58257	5.6	36	13
P1	gcvP, 2151957-2153414 (Clockwise) Hypothetical Protein gcvP (glycine cleavage system P protein)	sta_BASYS02061	96	53865	5.55	38	17
P2	glgA, 1304868-1306310 (Clockwise) Hypothetical Protein glgA (glgA, 1304868-1306310 (Clockwise) Hypothetical Protein glgA)	sta_BASYS01234	96	55873	5.6	34	14
P4	ypdF, 2851729-2849936 (CounterClockwise) Hypothetical Protein ypdF	sta_BASYS02718	152	68494	5.11	54	25
	aspS, 3407936-3406149 (CounterClockwise) Hypothetical Protein aspS	sta_BASYS03216	60	67951	5.04	33	16
P5	metG, 139988-138051 (CounterClockwise) Hypothetical Protein metG (methionyl-tRNA synthetase)	sta_BASYS00146	65	74724	5.21	23	12
P6	thrS, 918695-920725 (Clockwise) Hypothetical Protein thrS (threonyl-tRNA synthetase)	sta_BASYS00884	121	78293	5.42	39	21
P7	ygbD, 2298860-2297229 (CounterClockwise) Hypothetical Protein ygbD	sta_BASYS02209	114	59060	5.49	34	16
P10	ptsI, 3427518-3425806 (CounterClockwise) Hypothetical Protein ptsI (phosphoenolpyruvate-protein phosphotransferase)	sta_BASYS03230	182	63296	4.81	52	25
P11	groL, 567992-569620 (Clockwise) Hypothetical Protein groL (60 kDa chaperonin)	sta_BASYS00550	142	57747	4.74	48	19
P11	ppaC, 724318-722720 (CounterClockwise)	sta_BASYS00711	86	59727	4.86	30	14

	Probable manganese-dependent inorganic pyrophosphatase						
Strain B-1							
	Protein	**ID no**	**A**	**B**	**C**	**D**	**E**
A1	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	289	64664	4.76	65	40
A2	ATP synthase alpha chain	630_gi|115252530|emb|CAJ70373.1|	214	54688	4.91	54	27
A2	putative aminoacyl-histidine dipeptidase	630_gi|115249724|emb|CAJ67541.1|	74	53053	4.89	35	15
A3	ppaC, 1051638-1053236 (Clockwise) Probable manganese-dependent inorganic pyrophosphatase	stb_BASYS01013	167	59668	4.89	46	23
A5	electron transfer flavoprotein alpha-subunit	630_gi|115249407|emb|CAJ67222.1|	203	37353	5.03	61	24
A6	(R)-2-hydroxyisocaproate dehydrogenase	630_gi|115249400|emb|CAJ67215.1|	191	36552	5.07	62	20
A7	electron transfer flavoprotein beta-subunit	630_gi|115249406|emb|CAJ67221.1|	197	28767	4.87	80	21
A8	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	67	64664	4.76	24	12
A9	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	93	64664	4.76	26	15
A10	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	353	64664	4.76	65	37
A11	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072	stb_BASYS03072	259	64664	4.76	59	29
A12	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405|emb|CAJ67220.1|	224	41009	6.41	72	33
B1	pyruvate-flavodoxin oxidoreductase	630_gi|115251733|emb|CAJ69568.1|	155	130112	5.39	23	21
B2	pyruvate-flavodoxin oxidoreductase	630_gi|115251733|emb|CAJ69568.1|	231	130112	5.39	38	35
B3	pyruvate-flavodoxin oxidoreductase	630_gi|115251733|emb|CAJ69568.1|	270	130112	5.39	39	40
B5	pyruvate-flavodoxin oxidoreductase	630_gi|115251733|emb|CAJ69568.1|	256	130112	5.39	36	35
B6	tyrB, 2681281-2680022 (CounterClockwise) Hypothetical Protein tyrB (putative aspartate aminotransferase)	stb_BASYS02592	70	47863	5.29	26	9
B9	ileS, 2959099-2955983 (CounterClockwise) Hypothetical Protein ileS (isoleucyl-tRNA synthetase)	stb_BASYS02845	117	120425	5.08	25	18
B11	putative alanyl-tRNA synthetase	630_gi|115250316|emb|CAJ68138.1|	74	99077	4.94	17	12
B12	putative nitric oxide reductase flavoprotein	630_gi|115250664|emb|CAJ68488.1|	178	94987	4.92	29	21
C1	pheT, 1093175-1095568 (Clockwise) Phenylalanyl-tRNA synthetase beta chain	stb_BASYS00688	122	89576	4.87	24	15
C2	translation elongation factor G	630_gi|115249074|emb|CAJ66885.1|	205	76076	4.91	47	24

C3	translation elongation factor G	630_gi|115249074|emb|CAJ66885.1|	210	76076	4.91	45 %	23
C6	chaperone protein	630_gi|115251515|emb|CAJ69348.1|	200	66541	4.72	52 %	20
C8	htpG, 312628-310691 (CounterClockwise) Hypothetical Protein htpG (chaperone protein (heat shock protein))	630_gi|115249282|emb|CAJ67095.1|	275	75476	4.91	60 %	33
C9	atpD, 3370184-3368406 (CounterClockwise) Hypothetical Protein atpD (V-type sodium ATP synthase subunit A)	stb_BASYS03178	266	65694	4.89	63 %	35
C10	glyS, 3040737-3038671 (CounterClockwise) Hypothetical Protein glyS (glycyl-tRNA synthetase beta chain)	stb_BASYS02649	173	78491	4.97	32 %	20
C11	pyruvate kinase	630_gi|115252454|emb|CAJ70297.1|	94	63470	5	32 %	14
D1	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249403|emb|CAJ67218.1|	177	47133	5	51 %	25
D2	elongation factor TU	630_gi|115249061|emb|CAJ66872.1|	227	44226	4.94	72 %	27
D3	glycine reductase complex component B alpha and beta subunits	630_gi|115251407|emb|CAJ69239.1|	194	46436	4.95	58 %	18
D4	rpsA, 1080759-1082033 (Clockwise) Hypothetical Protein rpsA (putative 30S ribosomal protein S1)	stb_BASYS01043	80	48454	5.02	38 %	10
D5	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872|emb|CAJ67689.1|	127	58649	5.43	35 %	17
D6	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872|emb|CAJ67689.1|	208	58649	5.43	55 %	29
D7	oligopeptide ABC transporter, substrate-binding protein	630_gi|115251723|emb|CAJ69558.1|	156	58299	5.36	46 %	20
D8	glycine/sarcosine/betaine reductase complex component C beta subunit	630_gi|115251404|emb|CAJ69236.1|	130	55464	5.27	56 %	23
D8	GMP synthase glutamine-hydrolyzing	630_gi|115249206|emb|CAJ67019.1|	105	57802	5.26	42 %	19
D9	glycine/sarcosine/betaine reductase complex component C beta subunit	630_gi|115251404|emb|CAJ69236.1|	131	55464	5.27	41 %	17
D10	aegA, 1712608-1714002 (Clockwise) Hypothetical Protein aegA (putative glutamate synthase NADPH small chain)	stb_BASYS01667	229	50234	5.12	61 %	25
D11	4-aminobutyrate aminotransferase	630_gi|115251211|emb|CAJ69042.1|	158	48558	5.17	39 %	14
D12	ATP synthase beta chain	630_gi|115252528|emb|CAJ70371.1|	251	49817	5.18	87 %	32
E1	putative amino acid aminotransferase	630_gi|115252729|emb|CAJ70573.1|	125	44997	5.12	50 %	18
E2	putative amino acid aminotransferase	630_gi|115252729|emb|CAJ70573.1|	170	44997	5.12	62 %	29
E3	NAD-specific glutamate dehydrogenase	630_gi|115249189|emb|CAJ67001.1|	180	46271	5.16	66 %	26
E4	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	630_gi|115249401|emb|CAJ67216.1|	184	44491	5.18	73 %	29
E5	acetyl-CoA acetyltransferase	630_gi|115250080|emb|CAJ67900.1|	198	41005	5.2	71 %	24
E6	NAD-specific glutamate dehydrogenase	630_gi|115249189|emb|CAJ67001.1|	106	46271	5.16	44 %	18
E6	putative amino acid aminotransferase	630_gi|115252729|emb|CAJ70573.1|	88	44997	5.12	51 %	19
E6	aspartate aminotransferase	630_gi|115250375|	71	45066	5.1	38	15

		emb|CAJ68197.1|				%	
E7	hisC, 2820848-2821951 (Clockwise) Hypothetical Protein hisC (putative histidinol-phosphate aminotransferase)	stb_BASYS02725	153	42016	5.09	58 %	17
E8	putative tellurite resistance protein	630_gi|115250680| emb|CAJ68504.1|	212	43646	5.09	72 %	24
E9	elongation factor Ts	630_gi|115251193| emb|CAJ69024.1|	109	33290	5.21	41 %	12
E11	putative oxidoreductase, thiamine diP-binding subunit	630_gi|115249125| emb|CAJ66936.1|	118	39290	5.32	62 %	16
E12	butyrate kinase	630_gi|115249122| emb|CAJ66933.1|	197	38728	5.34	73 %	22
E12	putative oxidoreductase, thiamine diP-binding subunit	630_gi|115249125| emb|CAJ66936.1|	55	39290	5.32	48 %	11
F1	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249404| emb|CAJ67219.1|	200	42966	5.27	71 %	29
F2	electron transfer flavoprotein alpha-subunit	630_gi|115249407| emb|CAJ67222.1|	202	37353	5.03	55 %	17
F3	electron transfer flavoprotein alpha-subunit	630_gi|115250077| emb|CAJ67897.1|	167	36221	5.11	65 %	20
F4	thioredoxin reductase	630_gi|115251409| emb|CAJ69241.1|	128	34471	4.91	48 %	14
F5	ABC transporter, substrate-binding lipoprotein	630_gi|115249887| emb|CAJ67706.1|	210	36159	5.4	78 %	20
F6	flagellar hook protein	630_gi|115249263| emb|CAJ67076.1|	147	34568	5.24	56 %	14
F7	putative fructose-bisphosphate aldolase	630_gi|115249409| emb|CAJ67224.1|	220	33069	5.16	81 %	19
F8	putative O-acetylserine sulfhydrylase	630_gi|115250635| emb|CAJ68459.1|	143	32816	5.35	60 %	15
F10	nitrilase (carbon-nitrogen hydrolase)	630_gi|115251895| emb|CAJ69730.1|	62	34790	5.27	23 %	7
F11	ribose-phosphate pyrophosphokinase	630_gi|115252575| emb|CAJ70418.1|	107	34548	5.41	50 %	13
G1	electron transfer flavoprotein beta-subunit	630_gi|115249406| emb|CAJ67221.1|	157	28767	4.87	75 %	15
G2	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,methyltransferase subunit	630_gi|115249744| emb|CAJ67561.1|	76	29762	4.78	30 %	8
G3	fliY, 2449549-2448746 (CounterClockwise) Hypothetical Protein fliY (probable amino-acid ABC transporter,substrate-binding protein)	stb_BASYS02377	147	30120	5.22	53 %	13
G4	electron transfer flavoprotein beta-subunit	630_gi|115249406| emb|CAJ67221.1|	142	28767	4.87	61 %	13
G6	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,nickel-inserting subunit	630_gi|115249741| emb|CAJ67558.1|	89	28419	5.22	51 %	10
G7	GTP-sensing transcriptional pleiotropic repressor	630_gi|115250309| emb|CAJ68131.1|	126	28727	5.31	49 %	14
G9	putative sigma 54 modulation protein	630_gi|115251497| emb|CAJ69330.1|	61	21847	4.99	25 %	6
G10	glycine reductase complex component B alpha and beta subunits	630_gi|115251407| emb|CAJ69239.1|	158	46436	4.95	50 %	18
G12	putative subunit of oxidoreductase	630_gi|115249127| emb|CAJ66938.1|	112	19942	4.97	64 %	10
H2	putative dehydrogenase, electron transfer subunit	630_gi|115250577| emb|CAJ68401.1|	127	33353	5.49	65 %	18

H3	putative indolepyruvate oxidoreductase subunit	630_gi|115251432|emb|CAJ69265.1|	82	21020	5.44	52 %	7
H6	putative phosphosugar isomerase	630_gi|115252083|emb|CAJ69921.1|	50	21423	5.71	21 %	4
I2	adhE, 3383815-3381161 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase)	stb_BASYS03189	289	97214	5.63	54	41
I3	chaperone	630_gi|115251074|emb|CAJ68905.1|	103	97722	5.35	29	23
I6	formate acetyltransferase	630_gi|115249776|emb|CAJ67593.1|	76	83966	5.36	25	15
I7	pnp, 1460101-1462203 (Clockwise) Hypothetical Protein pnp (polyribonucleotide nucleotidyltransferase)	stb_BASYS01433	75	77769	5.37	19	12
I8	peptidase	630_gi|115251315|emb|CAJ69146.1|	113	68523	5.14	37	16
I10	threonyl-tRNA synthetase	630_gi|115249589|emb|CAJ67406.1|	67	73783	5.3	24	12
I11	indolepyruvate oxidoreductase	630_gi|115251433|emb|CAJ69266.1|	307	66900	5.72	61	33
I12	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252466|emb|CAJ70309.1|	112	69458	5.25	37	18
J4	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	630_gi|115249733|emb|CAJ67550.1|	74	69709	5.5	34	18
J5	putative translation elongation factor	630_gi|115249025|emb|CAJ66836.1|	61	72577	5.45	32	14
J6	putative translation elongation factor	630_gi|115249025|emb|CAJ66836.1|	171	72577	5.45	46	27
J7	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184|emb|CAJ66996.1|	228	69298	5.41	49	32
J8	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184|emb|CAJ66996.1|	91	69298	5.41	24	12
J11	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252466|emb|CAJ70309.1|	76	69458	5.25	24	14
K1	formate--tetrahydrofolate ligase	630_gi|115249735|emb|CAJ67552.1|	206	60347	5.61	55	25
K2	formate--tetrahydrofolate ligase	630_gi|115249735|emb|CAJ67552.1|	92	60347	5.61	38	17
K3	radical SAM-superfamily protein	630_gi|115251209|emb|CAJ69040.1|	216	52653	5.56	59	29
K4	succinate-semialdehyde dehydrogenase NAD(P)+	630_gi|115251397|emb|CAJ69229.1|	157	51163	5.56	50	21
K5	radical SAM-superfamily protein	630_gi|115251209|emb|CAJ69040.1|	137	15653	5.56	51	21
K6	glucose-6-phosphate isomerase	630_gi|115252341|emb|CAJ70182.1|	102	50677	5.48	33	13
K7	tyrosyl-tRNA synthetase	630_gi|115250562|emb|CAJ68386.1|	194	45703	5.42	51	22
K8	threonine dehydratase catabolic	630_gi|115251567|emb|CAJ69400.1|	139	43244	5.34	50	20

K9	acetate kinase	630_gi|115250207|emb|CAJ68028.1|	124	43691	5.4	53	18
K10	alanine racemase	630_gi|115252523|emb|CAJ70366.1|	191	43568	5.41	64	21
K11	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, alpha subunit	630_gi|115249743|emb|CAJ67560.1|	146	50477	5.32	37	16
K12	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	630_gi|115249053|emb|CAJ66864.1|	112	55737	5.29	43	19
K12	inosine-5'-monophosphate dehydrogenase	630_gi|115251390|emb|CAJ69222.1|	80	55085	5.32	37	15
L1	gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase	630_gi|115251396|emb|CAJ69228.1|	85	55864	5.8	47	16
L2	gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase	630_gi|115251396|emb|CAJ69228.1|	109	55864	5.8	33	12
L5	putative serine hydroxymethyltransferase	630_gi|115251778|emb|CAJ69613.1|	148	46051	5.67	48	20
L7	iscS, 1418080-1419288 (Clockwise) Hypothetical Protein iscS (cysteine desulfurase)	stb_BASYS01389	113	44510	5.9	39	14
L8	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405|emb|CAJ67220.1|	128	41009	6.41	45	17
L9	caiA, 1012108-1010963 (CounterClockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	stb_BASYS00979	56	41469	5.88	29	12
L10	cyclopropane-fatty-acyl-phospholipid synthase	630_gi|115249187|emb|CAJ66999.1|	108	46459	5.9	40	16
	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405|emb|CAJ67220.1|	89	41009	6.41	31	12
L11	butyryl-CoA dehydrogenase	630_gi|115250075|emb|CAJ67895.1|	185	41390	5.71	72	23
L12	glyceraldehyde-3-phosphate dehydrogenase 2	630_gi|115252231|emb|CAJ70071.1|	183	36239	5.72	58	19
M1	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	137	34339	7.82	51	16
M2	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	94	34339	7.82	44	11
M3	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	110	34339	7.82	44	12
M4	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	84	34339	7.82	47	13
M5	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	70	34339	7.82	47	12
M6	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	133	34339	7.82	47	13
M7	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	126	34339	7.82	55	17
M8	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	114	34339	7.82	51	17
M9	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	99	34339	7.82	47	17

M10	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	158	34339	7.82	57	21
M11	phosphate butyryltransferase	630_gi|115249121|emb|CAJ66932.1|	134	32670	5.61	63	19
	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC (flagellin subunit)	stb_BASYS00368	107	34339	7.82	52	15
M12	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079|emb|CAJ67899.1|	148	30887	5.8	56	14
N1	putative subunit of oxidoreductase	630_gi|115249126|emb|CAJ66937.1|	130	27031	6.38	59	10
N2	fliC, 349752-348793 (CounterClockwise) Hypothetical Protein fliC	stb_BASYS00368	87	34339	7.82	42	11
N3	30S ribosomal protein S2	630_gi|115251194|emb|CAJ69025.1|	93	26940	6.19	50	12
N4	activator of 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249402|emb|CAJ67217.1|	120	28948	6.35	43	13
N7	BASYS03019, 3175744-3173333 (CounterClockwise) Hypothetical Protein BASYS03019 (cell surface protein (putative cell surface-associated cysteine protease))	stb_BASYS03019	152	87454	7.49	34	21
N8	ATP-dependent Clp protease	630_gi|115249029|emb|CAJ66840.1|	88	91292	8.85	28	20
Strain Tra5/5							
	Protein	**ID no**	**A**	**B**	**C**	**D**	**E**
A1	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898	stt_BASYS02898	57	80379	4.83	34	18
A2	atpA, 3871191-3869689 (CounterClockwise) ATP synthase subunit alpha	stt_BASYS03613	154	54688	4.91	48	23
A3	ppaC, 409388-407790 (CounterClockwise) Probable manganese-dependent inorganic pyrophosphatase	stt_BASYS00431	70	59668	4.89	31	10
A5	fixB, 450899-451936 (Clockwise) Hypothetical Protein fixB (electron transfer flavoprotein alpha-subunit)	stt_BASYS00466	98	57353	5.03	44	14
A6	ldhA, 443798-442800 (CounterClockwise) Hypothetical Protein ldhA (R)-2-hydroxyisocaproate dehydrogenase)	stt_BASYS00458	98	36552	5.07	41	12
A9	tig, 3643244-3641901 (CounterClockwise) Hypothetical Protein tig (trigger factor)	stt_BASYS03407	82	50540	4.61	35	11
A11	caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	stt_BASYS00464	113	41469	5.88	46	17
A12	caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA	stt_BASYS00464	69	41469	5.88	39	12
B1	pyc, 12964-16395 (Clockwise) Pyruvate carboxylase	stt_BASYS00016	92	129243	5.5	17	18
B2	ydbK, 2906795-2903256 (CounterClockwise) Hypothetical Protein ydbK	stt_BASYS02779	113	130118	5.39	32	34
B3	ydbK, 2906795-2903256 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase0	stt_BASYS02779	143	130118	5.39	29	33
B4	ydbK, 2906795-2903256 (CounterClockwise) Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase1	stt_BASYS02779	108	130118	5.39	27	27
B5	ydbK, 2906795-2903256 (CounterClockwise)	stt_BASYS02779	87	130118	5.39	28	21

	Hypothetical Protein ydbK (pyruvate-flavodoxin oxidoreductase2						
B6	tyrB, 2547350-2546091 (CounterClockwise) Hypothetical Protein tyrB (putative aspartate aminotransferase)	stt_BASYS02457	99	47849	5.29	28	13
B8	carB, 3998253-3995047 (CounterClockwise) Hypothetical Protein carB	stt_BASYS03737	95	120071	4.98	17	16
B9	ileS, 2826892-2823776 (CounterClockwise) Hypothetical Protein ileS (isoleucyl-tRNA synthetase)	stt_BASYS02709	208	120833	5.08	42	35
B12	norV, 1741635-1744166 (Clockwise) Hypothetical Protein norV (putative nitric oxide reductase flavoprotein)	stt_BASYS01695	122	94959	4.92	34	18
C1	pheT, 835717-838110 (Clockwise) Phenylalanyl-tRNA synthetase beta chain	stt_BASYS00839	164	89576	4.87	56	34
C2	fusA, 77070-79136 (Clockwise) Elongation factor G	stt_BASYS00069	94	76076	4.91	31	15
C3	fusA, 77070-79136 (Clockwise) Elongation factor G	stt_BASYS00069	174	76076	4.91	51	27
C4	BASYS00871, 876652-878778 (Clockwise) Hypothetical Protein BASYS00871	stt_BASYS00871	172	77451	4.69	35	27
C5	infB, 1392980-1394920 (Clockwise) Translation initiation factor IF-2	stt_BASYS01361	78	69895	4.76	40	13
C6	dnaK, 2642740-2640893 (CounterClockwise) Chaperone protein dnaK	stt_BASYS02546	213	66541	4.72	57	26
C7	groL, 204704-206332 (Clockwise) 60 kDa chaperonin (60 kDa chaperonin)	stt_BASYS00202	230	57735	4.74	62	27
C8	htpG, 289359-291296 (Clockwise) Hypothetical Protein htpG (chaperone protein (heat shock protein))	stt_BASYS00289	227	75476	4.91	43	25
C9	atpA, 3180789-3179011 (CounterClockwise) V-type ATP synthase alpha chain	stt_BASYS03006	155	65712	4.89	46	23
C10	glyS, 2611727-2609661 (CounterClockwise) Hypothetical Protein glyS (glycyl-tRNA synthetase beta chain)	stt_BASYS02514	95	78493	4.97	32	15
C11	glyS, 2611727-2609661 (CounterClockwise) Hypothetical Protein glyS (glycyl-tRNA synthetase beta chain)	stt_BASYS02514	208	78493	4.97	43	27
D1	BASYS00462, 446522-447754 (Clockwise) Hypothetical Protein BASYS00462 (subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase)	stt_BASYS00462	104	471331 40	5	40	21
D2	elongation factor TU	630_gi|115249061|emb|CAJ66872.1|	123	44226	4.94	52	19
D3	rpsA, 1074685-1075959 (Clockwise) Hypothetical Protein rpsA (putative 30S ribosomal protein S1)	stt_BASYS01053	168	48455	4.98	59	21
D3	elongation factor TU	630_gi|115249061|emb|CAJ66872.1|	80	44226	4.94	39	13
D4	BASYS02894, 3042534-3040699 (CounterClockwise) Hypothetical Protein BASYS02894	stt_BASYS02894	75	67041	5.14	30	13
D4	aegA, 1647197-1648591 (Clockwise) Hypothetical Protein aegA (putative glutamate synthase NADPH small chain)	stt_BASYS01605	154	50248	5.12	52	18

D5	yddS, 2891711-2893306 (Clockwise) Hypothetical Protein yddS (oligopeptide ABC transporter, substrate-binding protein)	stt_BASYS02768	98	58299	5.51	24	11
D6	aegA, 1647197-1648591 (Clockwise) Hypothetical Protein aegA (putative glutamate synthase NADPH small chain)	stt_BASYS01605	100	50248	5.12	30	14
D7	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	630_gi|115249053|emb|CAJ66864.1|	92	49417	5.22	33	16
D8	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	630_gi|115249740|emb|CAJ67557.1|	83	50061	5.27	34	14
D8	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	630_gi|115249740|emb|CAJ67557.1|	91	50061	5.27	39	14
D9	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	630_gi|115249053|emb|CAJ66864.1|	74	59417	5.22	34	17
D9	atpB, 3179014-3177641 (CounterClockwise) V-type ATP synthase beta chain	stt_BASYS03005	105	50585	5.21	39	12
D10	atpB, 3179014-3177641 (CounterClockwise) V-type ATP synthase beta chain (V-type sodium ATP synthase subunit B)	stt_BASYS03005	149	50585	5.21	68	22
D11	atpD, 3868770-3867376 (CounterClockwise) ATP synthase subunit beta (ATP synthase beta chain)	630_gi|115252528|emb|CAJ70371.1|	176	49817	5.18	79	28
D12	atpB, 3179014-3177641 (CounterClockwise) V-type ATP synthase beta chain	630_gi|115252011|emb|CAJ69847.1|	109	50585	5.21	61	21
D12	elongation factor TU	630_gi|115249061|emb|CAJ66872.1|	80	44226	4.94	39	13
E1	yjiR, 4076689-4075493 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	stt_BASYS03814	75	44997	5.12	47	17
E2	yjiR, 4076689-4075493 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	stt_BASYS03814	138	44997	5.12	61	26
E2	gluD, 189677-190942 (Clockwise) NAD-specific glutamate dehydrogenase	stt_BASYS00188	60	46271	5.16	45	15
E3	gluD, 189677-190942 (Clockwise) NAD-specific glutamate dehydrogenase	stt_BASYS00188	113	46271	5.16	57	22
E4	frc, 444475-445674 (Clockwise) Hypothetical Protein frc (isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase)	stt_BASYS00460	78	44491	5.18	40	12
E5	atoB, 1171926-1173158 (Clockwise) Hypothetical Protein atoB (acetyl-CoA acetyltransferase)	stt_BASYS01130	92	43350	5.56	54	17
E6	yjiR, 302052-300856 (CounterClockwise) Hypothetical Protein yjiR (putative amino acid aminotransferase)	stt_BASYS03814	137	44997	5.12	58	26
E7	yfbQ, 3096018-3097217 (Clockwise) Hypothetical Protein yfbQ (putative aromatic amino acid aminotransferase)	stt_BASYS02934	189	45056	5.01	63	21
E8	hisC, 2688605-2689708 (Clockwise) Hypothetical Protein hisC (putative histidinol-phosphate aminotransferase)	stt_BASYS02589	96	42016	5.09	46	13
E9	tsf, 2273991-2273080 (CounterClockwise) Elongation factor Ts (elongation factor Ts)	stt_BASYS02200	96	33290	5.21	52	16

E10	BASYS00118, 108979-110058 (Clockwise) Hypothetical Protein BASYS00118 (putative oxidoreductase, thiamine diP-binding subunit)	stt_BASYS00118	119	39290	5.32	71	23
E11	BASYS00118, 108979-110058 (Clockwise) Hypothetical Protein BASYS00118 (putative oxidoreductase, thiamine diP-binding subunit)	stt_BASYS00118	88	39290	5.32	65	16
E12	ackA, 106946-108025 (Clockwise) Hypothetical Protein ackA (butyrate kinase)	stt_BASYS00115	166	38728	5.34	68	25
F1	atoB, 1171926-1173158 (Clockwise) Hypothetical Protein atoB (acetyl-CoA acetyltransferase)	stt_BASYS01130	69	43350	5.56	39	14
F2	fixB, 450899-451936 (Clockwise) Hypothetical Protein fixB (electron transfer flavoprotein alpha-subunit)	stt_BASYS00466	153	37353	5.03	43	17
F3	ydiR, 1169127-1170137 (Clockwise) Hypothetical Protein ydiR (electron transfer flavoprotein alpha-subunit)	stt_BASYS01127	101	36221	5.11	48	13
F5	ldhA, 443798-442800 (CounterClockwise) Hypothetical Protein ldhA ((R)-2-hydroxyisocaproate dehydrogenase)	stt_BASYS00458	99	36552	5.07	40	12
F6	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898	stt_BASYS02898	88	79874	4.79	35	21
F7	agaY, 455111-454185 (CounterClockwise) Hypothetical Protein agaY (putative fructose-bisphosphate aldolase)	stt_BASYS00468	90	33053	5.16	78	15
F8	cysM, 1712319-1713227 (Clockwise) Hypothetical Protein cysM (putative O-acetylserine sulfhydrylase)	stt_BASYS01663	148	32816	5.35	73	20
F8	BASYS03008, 3182193-3181216 (CounterClockwise) Hypothetical Protein BASYS03008 (V-type sodium ATP synthase subunit C)	stt_BASYS03008	70	37383	5.42	52	16
F9	cysM, 1712319-1713227 (Clockwise) Hypothetical Protein cysM (putative O-acetylserine sulfhydrylase)	stt_BASYS01663	132	32816	5.35	69	17
F10	galU, 2947127-2946183 (CounterClockwise) Hypothetical Protein galU	stt_BASYS02813	55	35957	6.5	33	9
F11	prs, 3916457-3915507 (CounterClockwise) Hypothetical Protein prs (ribose-phosphate pyrophosphokinase)	stt_BASYS03658	82	34548	5.41	45	9
F12	ycgM, 996287-997189 (Clockwise) Hypothetical Protein ycgM (putative hydrolase)	stt_BASYS00983	116	33967	5.07	75	15
G1	ydiQ, 450095-450880 (Clockwise) Hypothetical Protein ydiQ (electron transfer flavoprotein beta-subunit)	stt_BASYS00465	100	28767	4.87	57	12
G2	metH, 875789-876595 (Clockwise) Hypothetical Protein metH (putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,methyltransferase subunit)	stt_BASYS00870	79	29762	4.78	51	14
G3	ydiQ, 450095-450880 (Clockwise) Hypothetical Protein ydiQ (electron transfer flavoprotein beta-subunit)	stt_BASYS00465	147	28767	4.87	57	15
G4	paaF, 1170156-1170953 (Clockwise) Hypothetical Protein paaF (3-hydroxybutyryl-CoA dehydratase)	stt_BASYS01128	85	28628	4.97	53	10
G5	BASYS01173, 1213229-1214098 (Clockwise)	stt_BASYS01173	98	32880	5.24	56	12

	Hypothetical Protein BASYS01173 (nitroreductase-family protein)						
G6	BASYS00867, 872607-873377 (Clockwise) Hypothetical Protein BASYS00867 (putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,nickel-inserting subunit)	stt_BASYS00867	89	28419	5.22	48	11
G7	BASYS01321, 1356204-1356989 (Clockwise) Hypothetical Protein BASYS01321 (GTP-sensing transcriptional pleiotropic repressor)	stt_BASYS01321	79	28727	5.31	40	8
G8	BASYS01847, 1900865-1901566 (Clockwise) Cytoplasmic Protein (conserved hypothetical protein)	stt_BASYS01847	149	27008	5.16	82	19
G9	yhbH, 2621911-2621330 (CounterClockwise) Hypothetical Protein yhbH (putative sigma 54 modulation protein)	stt_BASYS02527	82	22538	5.01	64	8
G11	efp, 1322094-1322666 (Clockwise) Hypothetical Protein efp (elongation factor P)	stt_BASYS01289	66	21572	4.93	48	7
G12	BASYS00120, 110811-111368 (Clockwise) Hypothetical Protein BASYS00120 (putative subunit of oxidoreductase)	stt_BASYS00120	122	19942	4.97	91	13
H1	BASYS01604, 1646289-1647197 (Clockwise) Hypothetical Protein BASYS01604 (putative dehydrogenase, electron transfer subunit)	stt_BASYS01604	112	33943	5.29	61	16
H2	BASYS00862, 867899-868531 (Clockwise) Hypothetical Protein BASYS00862 (methenyltetrahydrofolate cyclohydrolase)	stt_BASYS00862	96	23091	5.33	53	11
H3	BASYS02455, 2544241-2543666 (CounterClockwise) Hypothetical Protein BASYS02455 (putative indolepyruvate oxidoreductase subunit)	stt_BASYS02455	87	21020	5.44	59	8
H4	tktA, 2483057-2482245 (CounterClockwise) Hypothetical Protein tktA (transketolase)	stt_BASYS02393	91	29182	5.54	42	8
H7	rplJ, 65691-66221 (Clockwise) Hypothetical Protein rplJ (50S ribosomal protein L10)	stt_BASYS00061	82	19581	6.2	51	8
H8	BASYS02780, 2907882-2907313 (CounterClockwise) Propanediol Utilization Protein (putative propanediol utilization protein)	stt_BASYS02780	86	20079	5.97	56	9
H11	BASYS01708, 1757727-1758305 (Clockwise) Hypothetical Protein BASYS01708 (tellurium resistance protein)	stt_BASYS01708	103	21633	4.56	53	9
I1	adhE, 3194458-3191768 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase)	stt_BASYS03018	150	98613	5.69	40	30
I2	adhE, 3194458-3191768 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase)	stt_BASYS03018	109	98613	5.69	30	21
I5	leuS, 2716913-2714493 (CounterClockwise) Hypothetical Protein leuS (leuS, 3140547-3138127 (CounterClockwise) Hypothetical Protein leuS)	stt_BASYS02608	133	92973	5.13	32	21

I6	pflB, 913824-911593 (CounterClockwise) Hypothetical Protein pflB (pflB, 913824-911593 (CounterClockwise) Hypothetical Protein pflB)	stt_BASYS00902	145	83966	5.39	46	29
I7	pflB, 913824-911593 (CounterClockwise) Hypothetical Protein pflB (pflB, 913824-911593 (CounterClockwise) Hypothetical Protein pflB)	stt_BASYS00902	209	83966	5.39	57	37
I8	ypdF, 2419026-2417233 (CounterClockwise) Hypothetical Protein ypdF (peptidase)	stt_BASYS02329	93	68523	5.14	39	13
I9	thrS, 662786-664816 (Clockwise) Hypothetical Protein thrS (threonyl-tRNA synthetase)	stt_BASYS00693	116	78305	5.42	47	28
I10	thrS, 662786-664816 (Clockwise) Hypothetical Protein thrS (threonyl-tRNA synthetase)	stt_BASYS00693	84	78305	5.42	29	15
I11	nuoF, 3794945-3796831 (Clockwise) Hypothetical Protein nuoF (putative iron-only hydrogenase,electron-transferring subunit)	stt_BASYS03544	99	69458	5.25	33	18
I12	BASYS02815, 2949551-2947560 (CounterClockwise) Hypothetical Protein BASYS02815	stt_BASYS02815	112	78379	5.33	48	27
J3	BASYS00858, 862861-864780 (Clockwise) Hypothetical Protein BASYS00858 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	stt_BASYS00858	99	69693	5.5	36	20
J4	BASYS00858, 862861-864780 (Clockwise) Hypothetical Protein BASYS00858 (putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase)	stt_BASYS00858	123	69693	5.5	42	28
J5	fusA, 16625-18565 (Clockwise) Hypothetical Protein fusA (putative translation elongation factor)	stt_BASYS00017	221	72577	5.45	64	41
J6	fusA, 16625-18565 (Clockwise) Hypothetical Protein fusA (putative translation elongation factor)	stt_BASYS00017	192	72577	5.45	57	36
J7	BASYS00183, 183555-185456 (Clockwise) Hypothetical Protein BASYS00183 (putative oxidoreductase, acetyl-CoA synthase subunit)	stt_BASYS00183	75	69298	5.41	27	16
J8	BASYS00183, 183555-185456 (Clockwise) Hypothetical Protein BASYS00183 (putative oxidoreductase, acetyl-CoA synthase subunit)	stt_BASYS00183	110	69298	5.41	33	20
J9	glmS, 113151-114983 (Clockwise) Hypothetical Protein glmS (glucosamine--fructose-6-phosphate aminotransferase isomerizing)	stt_BASYS00122	196	67613	5.2	52	26
J11	BASYS02456, 2546072-2544243 (CounterClockwise) Hypothetical Protein BASYS02456 (indolepyruvate oxidoreductase subunit)	stt_BASYS02456	102	68911	5.65	43	22
J12	BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (cell surface protein (putative S-layer protein precursor))	stt_BASYS02896	178	66363	6.47	57	30
K1	BASYS00861, 866100-867794 (Clockwise) Hypothetical Protein BASYS00861 (formate--tetrahydrofolate ligase)	stt_BASYS00861	132	60894	5.53	48	24
K2	BASYS00861, 866100-867794 (Clockwise) Hypothetical Protein BASYS00861	stt_BASYS00861	129	60894	5.53	49	23

K3	thiH, 2292548-2293924 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	stt_BASYS02217	55	52653	5.56	44	15
K4	gcvP, 1787909-1789366 (Clockwise) Hypothetical Protein gcvP (glycine cleavage system P protein)	stt_BASYS01732	92	53820	5.55	46	17
K5	gcvP, 1787909-1789366 (Clockwise) Hypothetical Protein gcvP	stt_BASYS01732	83	53820	5.55	47	17
K5	glgA, 1043461-1044903 (Clockwise) Hypothetical Protein glgA	stt_BASYS01026	73	55774	5.5	43	16
K8	tdcB, 2704099-2702858 (CounterClockwise) Threonine dehydratase catabolic (threonine dehydratase catabolic)	stt_BASYS02601	142	44317	5.2	53	19
K9	ackA, 1264632-1265828 (Clockwise) Acetate kinase (acetate kinase)	stt_BASYS01228	148	43691	5.4	64	20
K10	tyrS2, 1627797-1629005 (Clockwise) Tyrosyl-tRNA synthetase 2 (tyrosyl-tRNA synthetase)	stt_BASYS01589	193	45703	5.42	60	22
K11	lpdA, 871151-872551 (Clockwise) Hypothetical Protein lpdA (putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit)	stt_BASYS00866	109	50061	5.27	44	12
K12	BASYS00869, 874372-875739 (Clockwise) Hypothetical Protein BASYS00869 (putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, alpha subunit)	stt_BASYS00869	143	50477	5.32	49	17
L1	BASYS02414, 2507063-2505567 (CounterClockwise) Hypothetical Protein BASYS02414 (gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase)	stt_BASYS02414	94	55864	5.8	36	15
L2	BASYS02414, 2507063-2505567 (CounterClockwise) Hypothetical Protein BASYS02414 (gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase)	stt_BASYS02414	95	55864	5.8	38	14
L5	glyA, 2963788-2962544 (CounterClockwise) Hypothetical Protein glyA (putative serine hydroxymethyltransferase)	stt_BASYS02828	71	46065	5.67	31	9
L6	iscS, 1359720-1360928 (Clockwise) Hypothetical Protein iscS (cysteine desulfurase)	stt_BASYS01326	78	44510	5.9	23	9
L7	iscS, 1359720-1360928 (Clockwise) Hypothetical Protein iscS (cysteine desulfurase)	stt_BASYS01326	138	44510	5.9	58	24
L8	caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	stt_BASYS00464	149	41469	5.88	69	24
L9	caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	stt_BASYS00464	102	41469	5.88	50	17
L10	caiA, 448917-450062 (Clockwise) Hypothetical Protein caiA (acyl-CoA dehydrogenase, short-chain specific)	stt_BASYS00464	110	41469	5.88	59	19
L10	cfa, 188449-187262 (CounterClockwise) Hypothetical Protein cfa (cyclopropane-fatty-acyl-phospholipid synthase)	stt_BASYS00186	80	46517	5.9	47	18

L11	caiA, 1167131-1168267 (Clockwise) Hypothetical Protein caiA (butyryl-CoA dehydrogenase)	stt_BASYS01125	140	41390	5.71	55	15
L12	gapC, 3482040-3481033 (CounterClockwise) Hypothetical Protein gapC (239lyceraldehydes-3-phosphate dehydrogenase 2)	stt_BASYS03262	176	36239	5.72	57	19
M1	metF, 870169-871122 (Clockwise) Methylenetetrahydrofolate Reductase (putative methylenetetrahydrofolate reductase)	stt_BASYS00865	76	34785	8.9	41	8
M2	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	89	34339	7.82	41	12
M3	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	92	34339	7.82	41	10
M4	pta, 861343-862254 (Clockwise) Hypothetical Protein pta (putative phosphate butyryltransferase)	stt_BASYS00857	140	32850	5.76	62	20
M5	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	83	34339	7.82	41	10
M6	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	74	34339	7.82	41	10
M7	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	82	34339	7.82	40	8
M8	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	87	34339	7.82	51	13
M9	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	79	34339	7.82	44	11
M10	fliC, 252235-253194 (Clockwise) Hypothetical Protein fliC	stt_BASYS00250	68	34339	7.82	42	10
M11	pta, 105971-106924 (Clockwise) Hypothetical Protein pta (phosphate butyryltransferase)	stt_BASYS00114	91	35126	8.35	48	11
M12	pgi, 3620557-3619208 (CounterClockwise) Hypothetical Protein pgi	stt_BASYS03385	83	50663	5.47	41	13
N1	BASYS00119, 110058-110810 (Clockwise) Hypothetical Protein BASYS00119 (putative subunit of oxidoreductase0	stt_BASYS00119	50	27031	6.38	38	9
N2	BASYS00119, 110058-110810 (Clockwise) Hypothetical Protein BASYS00119	stt_BASYS00119	77	27031	6.38	54	10
N3	rpsB, 2274783-2274070 (CounterClockwise) 30S ribosomal protein S2 (30S ribosomal protein S2)	stt_BASYS02201	85	26940	6.19	56	11
N4	yjiL, 445704-446513 (Clockwise) Hypothetical Protein yjiL (activator of 2-hydroxyisocaproyl-CoA dehydratase)	stt_BASYS00461	118	28948	6.35	51	11
N5	paaH, 1171051-1171896 (Clockwise) Hypothetical Protein paaH (3-hydroxybutyryl-CoA dehydrogenase)	stt_BASYS01129	122	30887	5.8	46	14
N6	paaH, 1171051-1171896 (Clockwise) Hypothetical Protein paaH	stt_BASYS01129	79	30887	5.8	43	9
N7	rnfC, 1226332-1227669 (Clockwise) Hypothetical Protein rnfC (electron transport complex protein)	stt_BASYS01187	82	48962	6.3	38	13
N8	clpB, 21011-23458 (Clockwise) Hypothetical Protein clpB (ATP-dependent Clp protease)	stt_BASYS00021	188	91292	5.83	48	36
N9	aroA, 1980984-1981985 (Clockwise) Hypothetical Protein aroA (3-phosphoshikimate 1-carboxyvinyltransferase)	stt_BASYS01918	99	38093	7.51	42	12

N10	yjeK, 2408425-2407157 (CounterClockwise) Hypothetical Protein yjeK (L-lysine 2,3-aminomutase)	stt_BASYS02321	92	49419	6.08	37	16
N11	kdsA, 1978885-1979898 (Clockwise) Hypothetical Protein kdsA (v0	stt_BASYS01916	117	36868	6.15	59	17
N12	BASYS00868, 873397-874341 (Clockwise) Hypothetical Protein BASYS00868 (putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, small subunit)	stt_BASYS00868	127	34041	4.77	75	16
R1	BASYS00860, 865090-865869 (Clockwise) Hypothetical Protein BASYS00860 (putative carbon monoxide dehydrogenase accessory protein)	stt_BASYS00860	79	28728	5.69	50	11
R3	ydiA, 2586352-2585510 (CounterClockwise) Putative phosphotransferase CD2411 (conserved hypothetical protein)	stt_BASYS02492	136	32006	5.95	74	23
R5	ydiB, 1984601-1985410 (Clockwise) Hypothetical Protein ydiB (shikimate dehydrogenase)	stt_BASYS01922	135	30246	6.34	59	12
R6	spo0A, 1301769-1302593 (Clockwise) Stage 0 sporulation protein A homolog	stt_BASYS01268	104	30888	6.33	64	12
R8	paaH, 1171051-1171896 (Clockwise) Hypothetical Protein paaH (3-hydroxybutyryl-CoA dehydrogenase)	630_gi|115250247|emb|CAJ68068.1|	96	30887	5.8	43	11
R9	caiA, 1167131-1168267 (Clockwise) Hypothetical Protein caiA (butyryl-CoA dehydrogenase)	stt_BASYS01125	150	41390	5.71	72	21
R10	ATIC, 997288-998463 (Clockwise) Bifunctional purine biosynthesis protein PURH (putative formyltransferase)	stt_BASYS00984	149	43870	9.99	100	2
R11	serS, 1987-3258 (Clockwise) Hypothetical Protein serS (v0	stt_BASYS00003	101	48781	5.68	43	15
R12	ygfH, 2505133-2503826 (CounterClockwise) Hypothetical Protein ygfH (4-hydroxybutyrate CoA transferase)	stt_BASYS02412	116	48257	5.6	50	20
P1	BASYS00861, 866100-867794 (Clockwise) Hypothetical Protein BASYS00861 (formate--tetrahydrofolate ligase)	stt_BASYS00861	115	60894	5.53	41	19
P2	folD, 868623-869495 (Clockwise) Bifunctional protein folD (putative FolD bifunctional protein includes: methylenetetrahydrofolate dehydrogenase; methenyltetrahydrofolate cyclohydrolase)	stt_BASYS00863	81	31610	5.67	36	10
P4	BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896 (v0	stt_BASYS02896	69	66363	6.47	28	12
P6	secA, 3048012-3045667 (CounterClockwise) Hypothetical Protein secA (preprotein translocase SecA subunit)	stt_BASYS02897	120	89352	5.68	31	22
P10	lysS, 3955558-3954029 (CounterClockwise) Lysyl-tRNA synthetase (lysyl-tRNA synthetase)	stt_BASYS03698	50	58539	5.12	28	10
P12	ackA, 2543613-2542543 (CounterClockwise) Hypothetical Protein ackA (butyrate kinase)	stt_BASYS02454	86	38869	5.29	35	10

Table A2 – *Proteins identified by 1D GE followed by LC-MS/MS analysis with identification number, predicted mass and number of unique peptides detected for each strain. The number of proteins of different types are show as follows; metabolic (M), surface and virulence (s/v), hypothetical (h), regulatory (R), transporter (t) and ribosomal (r).*

	Identification	Number	Mass (kDa)	No. unique peptides					
				A		B-1		Tra5/5	
				1	2	1	2	1	2
	All strains – 332 proteins	M = 207, s/v = 18, h = 24, R = 42, t= 12, r = 29							
2	electron transfer flavoprotein beta-subunit	630_gi|115249406	29	11	10	7	7	8	7
	pyruvate-flavodoxin oxidoreductase	630_gi|115251733	129	28	21	22	19	17	19
3	ydbK, 3033598-3030059 (CounterClockwise) Hypothetical Protein ydbK	stb_BASYS02909							
656	pyruvate-flavodoxin oxidoreductase	(115251733)	129	0	0	0	2	0	0
	BASYS00732, 727813-719681 (CounterClockwise) Hypothetical Protein BASYS00732	stb_BASYS00732 (115249677)	308	41	16	34	39	0	3
4	(toxin A) BASYS00995, 1036085-1040797 (Clockwise) Hypothetical Protein BASYS00995 (toxin	sta_BASYS00995							
205	A)	(115249677)	178	4	3	0	1	0	0
5	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405	41	13	10	11	13	9	8
6	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC	sta_BASYS00598	34	19	11	17	15	5	4
7	elongation factor TU	630_gi|115249061	44	12	8	11	11	7	8
8	ABC transporter, substrate-binding lipoprotein	630_gi|115249887	36	15	11	12	10	3	4
9	putative translation elongation factor	630_gi|115249025	72	16	14	16	16	13	12
10	putative amino acid aminotransferase	630_gi|115252729	45	15	8	10	10	7	8
11	electron transfer flavoprotein alpha-subunit	630_gi|115249407	37	14	11	12	9	6	5
12	putative rubrerythrin	630_gi|115250565	20	12	4	5	4	3	3
13	BASYS01063, 1122044-1123738 (Clockwise) Hypothetical Protein BASYS01063 (formate--tetrahydrofolate ligase	sta_BASYS01063 (115249735)	61	14	10	10	13	12	11
14	activator of 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249402	29	9	9	9	9	9	8
16	DNA-directed RNA polymerase beta chain	630_gi|115249070	139	16	13	17	18	14	12
17	50S ribosomal protein L5	630_gi|115249089	20	11	8	10	10	5	11
18	acetyl-CoA acetyltransferase	630_gi|115250080	41	15	13	8	9	6	9
19	NAD-specific glutamate dehydrogenase	630_gi|115249189	46	10	7	11	8	10	8
20	putative oxidoreductase, thiamine diP-binding subunit	630_gi|115249125	39	8	6	4	5	3	5
21	pyc, 12964-16395 (Clockwise) Pyruvate carboxylase	stt_BASYS00016	129	18	12	10	13	12	11
22	ATP synthase alpha chain	630_gi|115252530	55	11	12	10	11	7	8

23	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079	31	11	6	7	7	5	7
24	groL, 567992-569620 (Clockwise) Hypothetical Protein groL	sta_BASYS00550	58	15	12	10	11	6	8
25	electron transfer flavoprotein beta-subunit	630_gi|115250076	29	13	8	10	10	9	9
26	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	630_gi|115249733	69	12	10	5	6	10	10
27	adhE, 3383815-3381161 (CounterClockwise) Hypothetical Protein adhE (aldehyde-alcohol dehydrogenase)	stb_BASYS03189 (115252023)	97	6	6	11	11	11	10
29	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249404	42	6	5	8	8	8	8
30	BASYS02833, 2973699-2971870 (CounterClockwise) Hypothetical Protein BASYS02833 (indolepyruvate oxidoreductase)	sta_BASYS02833 (115251433)	68	12	9	10	12	8	9
31	30S ribosomal protein S7	630_gi|115249073	18	9	5	9	6	6	5
32	30S ribosomal protein S10	630_gi|115249076	12	5	2	5	5	4	1
34	atpD, 3370184-3368406 (CounterClockwise) Hypothetical Protein atpD (V-type ATP synthase subunit A)	stb_BASYS03178 (115452012)	66	9	8	8	9	9	9
35	glyceraldehyde-3-phosphate dehydrogenase 2	630_gi|115252231	36	8	5	6	3	5	4
36	isoleucyl-tRNA synthetase	630_gi|115251669	120	16	8	7	8	8	8
37	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184	68	9	5	11	13	6	8
38	putative dehydrogenase, electron transfer subunit	630_gi|115250577	33	9	4	8	10	7	8
39 427	putative glutamate synthase NADPH small chain aegA, 1953842-1955236 (Clockwise) Hypothetical Protein aegA putative oxidoreductase	630_gi|115250578 sta_BASYS01860 (115250578)	50 50	6 2	5 2	9 0	8 0	8 0	8 0
40	cell surface protein (putative penicillin-binding protein)	630_gi|115250510	111	11	16	7	6	3	3
41	formate acetyltransferase	630_gi|115249776	84	8	9	10	9	10	8
42	30S ribosomal protein S2	630_gi|115251194	27	6	5	7	5	5	6
43	(R)-2-hydroxyisocaproate dehydrogenase	630_gi|115249400	37	11	6	7	3	3	3
44	chaperone protein	630_gi|115251515	66	4	3	3	3	4	2
45	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,methyltransferase subunit	630_gi|115249744	29	8	8	7	6	9	7
46	ATP synthase beta chain	630_gi|115252528	50	10	10	12	11	5	9
47	translation elongation factor G	630_gi|115249074	76	13	8	7	9	6	6
48	butyryl-CoA dehydrogenase	630_gi|115250075	41	8	7	5	4	7	4
49	putative calcium-transporting ATPase	630_gi|115250678	101	13	6	10	13	1	7
50	glycine reductase complex component B alpha and beta subunits	630_gi|115251407	46	7	4	10	10	1	1
51	flagellin subunit	630_gi|115249247	31	1	1	2	2	1	1
52	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872	59	7	6	8	11	4	4

53	BASYS03154, 3331959-3331390 (CounterClockwise) Propanediol Utilization Protein	sta_BASYS03154	20	8	5	4	6	2	3
54	rpsH, 87631-88029 (Clockwise) 30S ribosomal protein S8	stt_BASYS00086	15	5	5	4	5	3	1
55	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, alpha subunit	630_gi|115249743	50	5	3	4	4	5	5
56	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252654	119	8	5	12	12	8	7
57	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit	630_gi|115249745	77	9	8	3	4	8	7
	ptsG, 3311616-3309544 (counter clockwise) Hypothetical protein ptsG (PTS system, glucose specific IIBC component)	sta_BASYS03136 (115251718)	75	5	1	5	6	2	4
58 454	ptsG, 3012861-3010912 (CounterClockwise) Hypothetical Protein ptsG PTS system, glucose-specific IIbc component	stb_BASYS02893 (115251718)	70	0	0	1	1	0	0
60	proS2, 53161-54693 (Clockwise) Prolyl-tRNA synthetase 2	stt_BASYS00046	59	3	6	3	6	7	9
61	30S ribosomal protein S5	630_gi|115249094	18	6	3	6	5	1	5
62	50S ribosomal protein L4	630_gi|115249078	24	9	2	4	6	2	3
63	putative universal stress protein	630_gi|115249829	15	8	2	7	5	5	0
64	enolase	630_gi|115252227	46	6	6	4	6	4	4
65	BASYS02800, 2945591-2944029 (CounterClockwise) Hypothetical Protein BASYS02800 (glycine/sarcosine/betaine reductase complex component C beta subunit)	sta_BASYS02800 (115251404)	56	4	4	6	7	4	4
66	preprotein translocase SecA subunit	630_gi|115249152	102	8	8	3	6	9	8
67	GTP-sensing transcriptional pleiotropic repressor	630_gi|115250309	29	5	4	6	6	5	4
68	50S ribosomal protein L9	630_gi|115252722	17	7	4	7	5	1	3
70	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein	630_gi|115250698	92	11	4	2	4	8	9
71	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	630_gi|115249740	49	6	9	4	4	6	11
72	proline reductase subunit proprotein	630_gi|115252300	68	6	3	4	7	3	4
73	butyrate kinase	630_gi|115249122	39	7	8	7	5	3	3
74	cell surface protein	630_gi|115251820	78	8	5	6	8	1	1
75	oligopeptide ABC transporter, substrate-binding protein	630_gi|115251723	58	6	5	6	10	2	3
76	cysteine desulfurase	630_gi|115250313	44	6	2	2	5	4	4
77	putative iron-only hydrogenase, catalytic subunit	630_gi|115252467	65	6	8	6	6	5	5
78	DNA-directed RNA polymerase beta' chain	630_gi|115249071	130	9	6	6	7	5	4
79	putative subunit of oxidoreductase	630_gi|115249127	20	5	4	2	6	3	3
80	AsnC-family transcriptional regulator	630_gi|115252605	16	7	6	5	6	3	2
81	subunit of oxygen-sensitive 2-	630_gi|115249403	47	5	3	4	5	6	4

	hydroxyisocaproyl-CoA dehydratase								
82	putative methylenetetrahydrofolate reductase	630_gi|115249739	32	7	5	2	4	6	6
83	putative alanyl-tRNA synthetase	630_gi|115250316	98	6	3	5	6	6	6
84	ribose-phosphate pyrophosphokinase	630_gi|115252575	34	10	3	4	4	3	4
85	50S ribosomal protein L13	630_gi|115249112	16	4	1	3	2	1	2
86	frc, 444475-445674 (Clockwise) Hypothetical Protein frc (isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase)	stt_BASYS00460 (115249401)	44	8	3	4	4	4	2
87	putative aliphatic sulfonates ABC transporter,substrate-binding lipoprotein	630_gi|115250525	36	5	7	6	4	2	2
90	butyrate kinase	630_gi|115251431	39	9	5	5	5	4	3
91	putative O-acetylserine sulfhydrylase	630_gi|115250635	33	9	4	4	3	3	4
92	putative formate acetyltransferase	630_gi|115252338	89	8	4	0	1	4	5
94	threonine dehydratase catabolic	630_gi|115251567	43	5	4	2	2	6	6
95	50S ribosomal protein L10	630_gi|115249067	19	7	3	6	3	2	2
96	30S ribosomal protein S3	630_gi|115249083	30	4	2	4	4	3	3
97	stage 0 sporulation protein A	630_gi|115250247	31	6	2	5	3	2	3
98	50S ribosomal protein L21	630_gi|115250193	11	4	1	3	3	2	1
99	electron transfer flavoprotein alpha-subunit	630_gi|115250077	36	6	4	5	5	3	1
100	putative carbon starvation protein	630_gi|115251425	53	3	2	3	4	0	2
101 605	putative aminoacyl-histidine dipeptidase pepD, 849602-848151 (CounterClockwise) Hypothetical Protein pepD putative aminoacyl-histidine	630_gi|115249724 stt_BASYS00847 (115249724)	53 53	4 0	3 0	6 0	2 0	5 1	5 1
102	chaperone protein (heat shock protein)	630_gi|115249282	75	5	2	4	8	3	6
103	putative fructose-bisphosphate aldolase	630_gi|115249409	33	5	2	5	4	2	3
104	ATP-dependent Clp protease	630_gi|115249029	91	9	4	3	1	5	5
105	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,nickel-inserting subunit	630_gi|115249741	28	5	2	4	4	3	3
106	10 chaperonin	630_gi|115249203	10	5	2	5	6	2	0
107	tellurium resistance protein	630_gi|115250676	22	4	4	3	4	3	3
108	leucyl-tRNA synthetase	630_gi|115251575	92	4	2	1	0	7	9
109	putative membrane protein	630_gi|115249895	39	2	2	3	3	2	3
110	cell surface protein	630_gi|115251837	73	9	7	5	3	1	3
111	DNA-binding protein HU	630_gi|115252557	10	2	3	3	2	3	0
112	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, small subunit	630_gi|115249742	34	9	3	4	4	2	0
113	norV, 2106146-2108677 (Clockwise) Hypothetical Protein norV (putative nitric oxide reductase flavoprotein)	sta_BASYS02024 (115250664)	94	3	3	6	6	4	2
114	ATP-dependent Clp protease proteolytic subunit	630_gi|115252361	21	5	3	3	4	1	2
115 478	BASYS03019, 3175744-3173333 (CounterClockwise) Hypothetical Protein BASYS03019 (cell surface protein (putative cell surface-associated cysteine protease) cell surface protein (putative cell surface-associated cysteine protease)	stb_BASYS03019 (115251840) 630_gi|115251840	87 87	5 2	3 2	4 0	7 0	0 0	1 0
116	30S ribosomal protein S4	630_gi|115249105	24	4	2	3	4	3	4
117	putative phage-related cell wall hydrolase	630_gi|115250945	29	3	3	4	3	1	2

	(endolysin)								
118	putative phosphate butyryltransferase	630_gi|115249732	33	8	3	5	4	2	3
119	putative tellurite resistance protein	630_gi|115250680	44	3	2	4	5	4	4
120	phosphate butyryltransferase	630_gi|115249121	32	6	3	3	3	2	3
121	(CounterClockwise) Hypothetical Protein rho (transcription termination factor Rho)	115252548	61	3	3	2	3	4	3
122	NifU-like protein	630_gi|115250314	16	4	4	4	3	2	1
123	50S ribosomal protein L15	630_gi|115249096	16	1	2	2	1	2	2
124	30S ribosomal protein S13	630_gi|115249103	14	1	1	3	2	2	0
126	putative histidinol-phosphate aminotransferase	630_gi|115251555	42	3	3	5	3	3	3
127	4-hydroxybutyrate CoA transferase	630_gi|115251394	48	5	5	3	3	2	2
128	putative ATP/GTP-binding protein	630_gi|115249123	25	5	2	4	5	3	2
129	30S ribosomal protein S9	630_gi|115249113	15	2	2	2	4	2	1
130	tig, 3643244-3641901 (CounterClockwise) Hypothetical Protein tig (trigger factor)	stt_BASYS03407 (115252362)	51	2	3	3	3	4	4
131	electron transport complex protein	630_gi|115250173	34	3	2	3	3	2	3
132	peptidyl-prolyl cis-trans isomerase	630_gi|115249340	19	3	3	3	3	2	3
133	putative transcriptional regulator	630_gi|115250535	23	2	2	2	2	2	2
134	putative RNA-binding protein	630_gi|115249151	21	5	2	3	4	2	1
135	BASYS00913, 941972-944095 (Clockwise) Hypothetical Protein BASYS00913 (hypothetical protein)	sta_BASYS00913 (115249604)	79	8	3	1	3	1	6
136	glucosamine--fructose-6-phosphate aminotransferase isomerizing	630_gi|115249129	67	4	2	2	3	3	3
138	pheT, 835717-838110 (Clockwise) Phenylalanyl-tRNA synthetase beta chain	stt_BASYS00839	89	2	2	2	4	5	5
139	rubrerythrin	630_gi|115249842	21	5	2	3	2	2	2
140	3-hydroxybutyryl-CoA dehydratase	630_gi|115250078	28	5	2	2	2	2	3
141	proline reductase	630_gi|115252298	26	4	2	4	4	2	2
142	putative carbon monoxide dehydrogenase accessory protein	630_gi|115249734	29	2	3	2	0	3	4
143	ppiC, 1744732-1743947 (CounterClockwise) Hypothetical Protein ppiC (putative peptidyl-prolyl isomerise)	sta_BASYS01657 (115250393)	30	2	0	8	7	1	0
144	dxs, 2909947-2908955 (CounterClockwise) Hypothetical Protein dxs (transketolase)	sta_BASYS02770 (115251376)	35	5	1	3	3	3	3
145	dihydrodipicolinate synthase	630_gi|115252280	32	6	2	4	4	2	1
146	anti-sigma F factor	630_gi|115249788	16	3	2	5	4	2	0
147	nitroreductase-family protein	630_gi|115250155	33	3	1	3	4	5	2
148	50S ribosomal protein L6	630_gi|115249092	20	2	2	2	4	1	2
149	V-type sodium ATP synthase subunit B	630_gi|115252011	51	4	3	4	3	1	3
150	succinate-semialdehyde dehydrogenase NAD(P)+	630_gi|115251397	51	2	3	2	0	3	4
152 803	clpB, 2575497-2572903 (CounterClockwise) Hypothetical Protein clpB (chaperone) chaperone	sta_BASYS02475 (115251074) 630_gi|115251074	98 98	4 0	2 0	3 0	3 1	4 0	5 0
153	DNA gyrase subunit A	630_gi|115249009	91	4	2	1	4	3	5
154	formylglycinamidine ribonucleotide synthetase	630_gi|115249233	141	7	4	3	4	1	1
155	shikimate kinase	630_gi|115250884	20	4	1	4	4	1	3
156	aspartate aminotransferase	630_gi|115250375	45	7	3	2	0	4	1

157	conserved hypothetical protein	630_gi|115252525	17	4	3	3	4	1	0
158	hydN, 446278-444749 (CounterClockwise) Hypothetical Protein hydN (Iron dependent hydrogenase)	stb_BASYS00490 (115249907)	56	0	1	2	5	5	4
159	thioredoxin reductase	630_gi|115251409	34	3	2	4	4	1	1
160	50S ribosomal protein L16	630_gi|115249084	16	2	2	2	2	0	2
161	gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase	630_gi|115251396	55	4	3	4	4	0	1
162	conserved hypothetical protein	630_gi|115251617	13	4	1	2	4	1	0
163	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252652	119	4	3	3	2	3	3
164	conserved hypothetical protein	630_gi|115250810	27	2	2	1	2	3	3
165	50S ribosomal protein L14	630_gi|115249087	13	3	1	3	3	2	0
169	manganese-dependent inorganic pyrophosphatase	630_gi|115249342	59	4	3	2	4	3	2
170	triosephosphate isomerase	630_gi|115252229	27	2	2	3	3	0	4
171	rubredoxin oxidoreductase (desulfoferrodoxin)	630_gi|115249844	14	3	2	2	2	2	0
172	BASYS00303, 304111-304506 (Clockwise) NimC/NimA Family Protein	stt_BASYS00303	15	3	1	1	2	1	0
173	conserved hypothetical protein	630_gi|115249812	26	1	2	2	2	3	1
174	electron transport complex protein	630_gi|115250168	48	2	1	1	1	3	2
175	cell surface protein	630_gi|115250068	37	4	5	1	3	3	1
176	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	630_gi|115252284	25	0	3	2	3	4	4
177	efp, 1620295-1620867 (Clockwise) Hypothetical Protein efp (Elongation factor P)	sta_BASYS01536 (115250279)	22	3	2	1	2	0	2
178	putative alanine racemase	630_gi|115251674	27	2	1	2	4	2	2
179	BASYS00661, 628949-629578 (Clockwise) Hypothetical Protein BASYS00661 (putative lipoprotein)	stt_BASYS00661 (115249561)	24	2	2	3	2	1	2
180	valS, 4034578-4031912 (CounterClockwise) Valyl-tRNA synthetase	sta_BASYS03755	103	2	3	0	1	3	3
181	lysyl-tRNA synthetase	630_gi|115252615	58	2	2	2	2	3	3
182	putative phosphatase	630_gi|115251616	60	3	2	2	2	2	1
183	yjiM, 1885289-1886440 (Clockwise) Hypothetical Protein yjiM (putative 2-hydroxyacyl-CoA dehydratase)	stt_BASYS01828 (115250793)	44	2	2	2	1	1	2
184	putative nitroreductase	630_gi|115251628	25	6	2	1	2	1	1
185	uracil phosphoribosyltransferase	630_gi|115252539	23	4	1	1	3	1	1
186	norV, 1245992-1244799 (CounterClockwise) Hypothetical Protein norV (putative anaerobic nitric oxide reductase flavorubredoxin)	stt_BASYS01207 (115250189)	45	1	1	1	1	1	1
187	putative cold shock protein	630_gi|115250391	7	1	1	1	1	1	1
189	ABC transporter, ATP-binding protein	630_gi|115249888	28	6	1	3	3	1	2
190	translation initiation factor IF-2	630_gi|115250345	70	5	1	1	4	1	1
191	conserved hypothetical protein	630_gi|115251463	32	4	0	4	4	2	2
192	rpsK, 94194-94592 (Clockwise) 30S ribosomal protein S11	stt_BASYS00099	14	2	2	3	2	2	0
193	ATP synthase B chain	630_gi|115252532	20	3	1	2	2	1	0
194	glycyl-tRNA synthetase beta chain	630_gi|115251485	78	5	1	2	2	2	2

195	aegA, 4105133-4104582 (CounterClockwise) Hypothetical Protein aegA (electron transport protein)	sta_BASYS03827 (115252371)	20	3	1	4	5	1	2
196	ribulose-phosphate 3-epimerase	630_gi|115251631	24	2	3	3	2	1	3
197	V-type sodium ATP synthase subunit E	630_gi|115252015	21	5	2	2	2	1	1
	BASYS01064, 1123843-1124475 (Clockwise) Hypothetical Protein BASYS01064 (methenyltetrahydrofolate cyclohydrolase)	sta_BASYS01064 (115249736)	23	8	1	1	1	1	1
199	V-type sodium ATP synthase subunit K	630_gi|115252016	16	2	1	1	1	1	1
200	peptidase	630_gi|115251315	68	3	3	2	2	2	0
201	putative indolepyruvate oxidoreductase subunit	630_gi|115251432	21	4	2	3	2	1	0
202	putative ATP-binding protein	630_gi|115250774	29	3	0	3	3	1	1
203	putative membrane protein	630_gi|115249537	42	1	1	1	1	1	1
204	ucpA, 407353-408141 (Clockwise) Hypothetical Protein ucpA (NADP-dependent 7-alpha-hydroxysteriod dehydrogenase	sta_BASYS00404 (115249069)	28	0	2	5	5	1	1
207	dihydrodipicolinate synthase	630_gi|115252282	32	3	2	3	2	2	3
208	LysR-family transcriptional regulator	630_gi|115250103	35	2	3	2	2	3	3
209	methionyl-tRNA synthetase	630_gi|115252601	74	4	2	1	2	2	1
210	elongation factor Ts	630_gi|115251193	33	5	2	0	1	1	2
211	putative membrane protein	630_gi|115250814	41	3	1	2	3	0	2
212	minD, 1239260-1240057 (Clockwise) Hypothetical Protein mind septum site-determining protein (cell division inhibitor)	stt_BASYS01200 (115250182)	29	3	1	3	3	1	2
213	rplA, 64794-65492 (Clockwise) Hypothetical Protein rplA (50S ribosomal protein L1)	stt_BASYS00060 (115249066)	25	2	1	2	2	3	1
214	putative ABC transporter, permease protein	630_gi|115250570	45	4	2	2	1	0	3
216	glycogen synthase	630_gi|115249898	56	2	2	2	2	2	2
217	BASYS02493, 2586997-2586359 (CounterClockwise) CBS Domain Protein (Conserved hypothetical protein)	stt_BASYS02493 (115251464)	23	2	2	2	2	1	2
218	GMP synthase glutamine-hydrolyzing	630_gi|115249206	57	4	2	1	3	1	1
219	polyribonucleotide nucleotidyltransferase	630_gi|115250354	78	4	2	1	2	1	1
220	pip, 3469683-3468751 (CounterClockwise) Proline iminopeptidase	stb_BASYS03264	36	2	1	3	3	3	1
221	ribose-5-phosphate isomerase 2	630_gi|115252540	16	2	2	2	2	3	0
223	acetate kinase	630_gi|115250207	43	3	2	1	1	2	2
224	nrdD, 449381-451732 (Clockwise) Hypothetical Protein nrdD (anaerobic ribonucleoside-triphosphate reductase)	sta_BASYS00451 (115249116)	89	2	2	1	4	1	2

225	sbcC, 1426372-1429920 (Clockwise) Hypothetical Protein sbcC (exonuclease subunit C)	sta_BASYS01330 (115250064)	137	3	2	2	3	1	3
226	ptsI, 2998875-2997163 (CounterClockwise) Hypothetical Protein ptsI (phosphoenolpyruvate-protein phosphotransferase)	stt_BASYS02857 (115251808)	63	1	1	1	1	1	3
227	metallo beta-lactamase superfamily protein	630_gi|115250323	62	3	1	0	3	2	3
228	heat shock protein	630_gi|115251516	24	2	0	4	3	1	3
229	shikimate dehydrogenase	630_gi|115250883	30	1	0	2	3	2	3
234	tyrosyl-tRNA synthetase	630_gi|115250562	46	2	2	2	1	3	2
235	putative cyclase	630_gi|115251016	24	4	1	3	2	1	2
236	fliY, 2754011-2753208 (CounterClockwise) Hypothetical Protein fliY (probable amino-acid ABC transporter, substrate-binding protein	sta_BASYS02640 (115251230)	30	2	2	2	2	1	1
237	probable amino-acid ABC transporter,substrate-binding protein	630_gi|115251227	29	1	1	2	4	2	4
238	ATP-dependent Clp protease ATP-binding subunit	630_gi|115252360	46	1	1	2	2	3	3
239	putative surface protein	630_gi|115251769	98	3	1	3	3	0	2
240	ATP synthase subunit gamma	630_gi|115252529	32	1	1	3	4	2	0
241	BASYS01536, 1574800-1575345 (Clockwise) Hypothetical Protein BASYS01536 (putative ruberythrin)	stt_BASYS01536 (115250515)	20	3	1	1	1	1	0
243	transcription antitermination protein	630_gi|115249064	20	3	1	3	2	2	2
244	30S ribosomal protein S19	630_gi|115249081	11	2	1	2	2	1	1
245	S-adenosylmethionine synthetase	630_gi|115249139	44	1	1	1	1	2	2
246	UTP--glucose-1-phosphate uridylyltransferase	630_gi|115251766	33	2	1	3	3	1	1
247	UDP-N-acetylenolpyruvoylglucosamine reductase	630_gi|115252462	33	1	1	2	2	1	1
248	pyruvate kinase	630_gi|115252454	63	3	1	1	2	1	1
249	cell-division initiation protein	630_gi|115251670	20	1	0	4	5	1	1
250	electron transfer flavoprotein beta-subunit	630_gi|115249821	29	1	1	1	1	1	1
251	ferritin	630_gi|115251248	20	1	1	1	1	1	1
252	putative membrane protein	630_gi|115249847	43	3	1	2	2	0	1
253	putative membrane protein	630_gi|115249032	40	3	1	1	2	0	1
256	BASYS00309, 303323-302691 (CounterClockwise) Hypothetical Protein PG (conserved hypothetical protein-selenium metabolism protein YedF	sta_BASYS00309 (115252732)	23	1	2	2	2	1	2
257	4-aminobutyrate aminotransferase	630_gi|115251211	48	3	1	2	2	1	1
258	transketolase, thiamine disphosphate-binding subunit	630_gi|115252520	30	1	1	3	1	1	1
259	excinuclease ABC subunit A	630_gi|115252471	105	3	4	0	1	2	2
260	putative cytosine permease	630_gi|115251790	45	2	1	3	3	0	1
261	fabG, 1329810-1330559 (Clockwise) Hypothetical Protein fabG (3-oxoacyl-[acyl-carrier protein] reductase	stb_BASYS01300 (115250215)	27	1	1	2	2	0	1
262	50S ribosomal protein L20	630_gi|115249703	13	2	0	3	2	2	0
263	putative DNA-binding protein	630_gi|115249180	23	4	1	2	1	0	1
264	BASYS00438, 424036-418979 (CounterClockwise) Hypothetical Protein	stt_BASYS00438 (115249349)	195	3	0	3	6	0	1

	BF (conserved hypothetical protein)								
266	DNA mismatch repair protein	630_gi|115251023	109	2	2	1	2	2	2
267	orotate phosphoribosyltransferase	630_gi|115249197	21	2	1	3	2	1	2
268	SpoIIIJ-associated protein	630_gi|115252741	24	2	1	2	2	1	1
269	putative imidazole glycerol phosphate synthase subunit	630_gi|115250592	23	2	1	2	2	1	1
270	putative aldo/keto reductase	630_gi|115249577	38	0	1	2	1	1	1
271	yjiL, 1886465-1887220 (Clockwise) Hypothetical Protein yjiL (putative (R)-2-hydroxyglutaryl-CoA dehydratase alpha subunit)	stt_BASYS01829 (115250794)	27	4	1	0	1	2	2
272	putative 30S ribosomal protein S1	630_gi|115250007	48	0	1	1	2	1	5
273	sigma-54-dependent transcriptional regulator	630_gi|115249408	63	0	1	3	3	1	0
274	threonyl-tRNA synthetase	630_gi|115249589	74	2	0	0	2	1	3
275	50S ribosomal protein L24	630_gi|115249088	11	1	1	1	1	1	0
276	50S ribosomal protein L3	630_gi|115249077	22	1	0	3	2	0	3
283	glycine reductase complex component B gamma subunit	630_gi|115251405	47	1	1	2	2	1	1
284	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252466	68	4	1	0	2	1	1
285	putative tRNA binding protein	630_gi|115251499	17	1	1	1	1	1	1
287	cell division protein	630_gi|115251697	41	1	1	1	1	1	1
288	yjiY, 2940632-2939211 (CounterClockwise) Hypothetical Protein yjiY (putative carbon starvation protein)	stb_BASYS02828 (115251657)	51	1	1	2	3	0	1
289	biotin carboxyl carrier protein of acetyl-CoA carboxylase	630_gi|115250985	17	2	1	1	1	0	1
290	pyruvate, phosphate dikinase	630_gi|115251462	97	1	0	0	1	4	3
296	DNA polymerase I	630_gi|115250159	101	2	1	1	1	1	1
297	uridylate kinase	630_gi|115251192	26	2	1	1	1	1	1
298	UDP-glucose 4-epimerase	630_gi|115251765	37	2	2	2	1	0	1
299	V-type sodium ATP synthase subunit D	630_gi|115252010	26	1	1	2	1	0	2
300	putative homocysteine S-methyltransferase	630_gi|115252660	87	0	1	2	1	3	4
301	putative carbonic anhydrase	630_gi|115251267	24	2	1	2	1	0	1
302	putative 1-(5-phosphoribosyl)-5-(5-phosphoribosylamino)methylidene amino imidazole-4-carboxamide isomerase	630_gi|115250593	27	2	1	1	1	1	0
303	ferredoxin	630_gi|115252670	6	1	1	1	1	1	0
93	ferredoxin	630_gi|115249124	8	3	5	5	1	0	0
727	BASYS01993, 2078571-2078759 (Clockwise) Hypothetical Protein BASYS01993 ferredoxin [*Clostridium difficile* QCD-23m63	sta_BASYS01993	7	1	0	0	0	0	0
308	two-component response regulator	630_gi|115251774	27	1	1	2	1	1	1
309	ABC transporter, ATP-binding protein	630_gi|115251122	62	1	1	1	0	3	1
310	adenylate kinase	630_gi|115249098	24	2	1	2	1	1	0
320 827	pepT, 4148458-4147277 (CounterClockwise) Hypothetical Protein pepT (putative peptidase) putative peptidase	stb_BASYS03876 (115252582) 630_gi|115252582	42 40	1 0	3 0	1 0	1 0	1 0	1 1
321	BASYS00366, 346535-344658 (CounterClockwise) Glycosyltransferase	stb_BASYS00366	72	1	1	1	2	1	1
322	rimM, 1332629-1333144 (Clockwise)	stt_BASYS01299	20	2	1	2	2	1	0

	Hypothetical Protein rimM (putative 16S rRNA-processing protein)	(115250289)							
323	NH3-dependent NAD(+) synthetase	630_gi|115249811	28	2	1	1	1	0	2
324	translation initiation factor IF-3	630_gi|115249701	19	1	2	0	2	1	2
325	cell division protein	630_gi|115252622	74	2	1	2	1	0	1
326	lon, 3638626-3636257 (CounterClockwise) Hypothetical Protein lon (ATP-dependent protease La)	stt_BASYS03401 (115252357)	89	1	1	0	1	1	2
327	putative NAD-dependent deacetylase (Sir2-family regulatory)	630_gi|115250338	27	1	0	2	1	1	1
328	putative NUDIX-family hydrolase	630_gi|115249807	17	1	1	1	1	1	0
329	imidazoleglycerol-phosphate dehydratase	630_gi|115250591	22	1	1	1	1	0	1
330	conserved hypothetical protein	630_gi|115250439	23	3	0	1	3	0	1
332	putative glycine cleavage system H protein	630_gi|115249746	14	0	2	2	1	2	0
333	putative membrane protein	630_gi|115251110	63	3	1	1	1	0	1
335	putative aldose epimerase	630_gi|115252500	34	1	0	2	2	1	0
340	putative exported carboxy-terminal processing protease	630_gi|115252515	42	1	0	1	0	1	0
346	UDP-N-acetylmuramoylalanine--D-glutamate ligase	630_gi|115251704	50	2	1	1	1	1	1
347	eutG, 2503854-2502688 (CounterClockwise) Hypothetical Protein eutG (putative propanediol/ethanolamine utilization protein	stt_BASYS02411 (115250967)	44	1	1	1	1	1	1
348	ABC transporter, ATP-binding protein	630_gi|115249802	59	1	1	2	1	0	1
349	conserved hypothetical protein	630_gi|115250191	27	1	1	1	2	1	1
350	hypothetical protein	630_gi|115251768	78	1	1	0	1	2	1
351	putative transaldolase	630_gi|115251384	23	1	1	1	1	0	1
353	conserved hypothetical protein	630_gi|115252075	17	1	1	1	1	1	0
354	putative translation inhibitor endoribonuclease	630_gi|115251566	14	2	0	1	1	1	0
359	DNA-directed RNA polymerase alpha chain	630_gi|115249106	35	5	0	1	0	1	0
364	phenylalanyl-tRNA synthetase alpha chain	630_gi|115249715	38	0	1	1	1	1	1
365	two-component system response regulator	630_gi|115251168	25	1	1	1	2	0	1
366	putative imidazole glycerol phosphate synthase subunit	630_gi|115250594	28	1	0	2	2	0	1
367	ribosome recycling factor	630_gi|115251191	21	2	1	1	1	1	1
369	pfkA, 3782241-3781282 (CounterClockwise) 6-phosphofructokinase	stt_BASYS03533	34	1	0	2	1	1	0
370	putative L-asparaginase	630_gi|115251569	36	0	1	1	2	3	0
372	BASYS01830, 1919083-1918436 (CounterClockwise) Hypothetical Protein CTC (hypothetical protein)	sta_BASYS01830 (115250551)	24	3	1	1	0	0	1
384	putative metal dependent phosphohydrolase	630_gi|115251577	22	1	1	1	1	0	1
385	putative inorganic polyphosphate/ATP-NAD kinase	630_gi|115250070	30	0	1	1	1	1	1
386	peptide deformylase 2	630_gi|115251641	16	1	1	1	1	1	0
387	putative transcriptional repressor	630_gi|115252578	32	0	2	1	2	0	2
388	hcr, 624539-622545 (CounterClockwise) Hypothetical Protein Mbar A	stb_BASYS00654 (115249747)	74	1	1	0	1	0	1

	(putative iron-sulfur protein)								
389	putative sigma 54 modulation protein	630_gi|115251497	22	1	0	1	1	1	0
391	inosine-5'-monophosphate dehydrogenase	630_gi|115251390	55	0	1	0	0	2	2
400	glycerol dehydratase	630_gi|115250149	89	0	1	2	1	1	1
401	putative phosphoglucomutase	630_gi|115250617	64	2	2	0	1	1	1
402	rplK, 64298-64723 (Clockwise) 50S ribosomal protein L11	stt_BASYS00059	15	1	1	1	1	1	0
406	ABC transporter, ATP-binding protein	630_gi|115252673	27	2	1	0	1	0	1
408	conserved hypothetical protein	630_gi|115252608	28	1	0	1	1	1	1
411	conserved hypothetical protein	630_gi|115250280	17	1	0	1	2	0	1
419	putative acetyltransferase	630_gi|115249853	19	1	0	2	0	0	1
431	putative membrane protein	630_gi|115250415	17	1	1	1	1	0	1
432	V-type sodium ATP synthase subunit I	630_gi|115252017	72	1	1	1	1	0	1
433	putative hydrolase	630_gi|115251794	23	1	1	0	1	1	0
434	glucose-6-phosphate isomerase	630_gi|115252341	51	1	0	1	0	1	2
435	putative signal recognition particle	630_gi|115250284	47	0	1	1	1	0	1
438	conserved hypothetical protein	630_gi|115250440	24	1	1	0	1	0	1
440	sigma-54-dependent transcriptional activator	630_gi|115252301	67	1	1	0	1	1	1
442	BASYS01208, 1246417-1248156 (Clockwise) Hypothetical Protein BASYS01208 (radical SAM-superfamily protein)	stt_BASYS01208 (115250190)	67	1	1	0	1	0	1
458	BASYS00601, 620689-622851 (Clockwise) Streptomycin Biosynthesis StrF Domain Protein	sta_BASYS00601	84	1	0	1	1	1	0
459	ribonuclease R	630_gi|115252220	82	1	1	1	0	1	0
460	cobalt-precorrin-3b c(17)-methyltransferase	630_gi|115252484	27	0	1	2	0	1	1
465	conserved hypothetical protein	630_gi|115251601	68	1	0	1	0	1	1
526	putative GTP cyclohydrolase I	630_gi|115250490	21	0	1	0	1	1	0
560	transcription-repair coupling factor	630_gi|115252562	130	1	0	0	1	1	0
574	PTS system, IIbc component	630_gi|115251321	68	1	0	1	0	0	1
615	BASYS00737, 756750-756154 (CounterClockwise) Deoxyribonuclease TatD Family	sta_BASYS00737	23	1	0	1	0	0	1
673	glycine betaine/carnitine/choline ABC transporter, ATP-binding protein	630_gi|115249914	42	0	1	0	0	1	0
	secA, 3048012-3045667 (CounterClockwise) Hypothetical Protein secA	stt_BASYS02897 (115251845)	89	3	2	0	0	1	3
292 206	(pre protein translocase secA) secA, 3237442-3235100 (CounterClockwise) Hypothetical Protein secA preprotein translocates secA subunit	stb_BASYS03071 (115251845)	90	0	0	5	6	0	0
	putative ornithine cyclodeaminase	630_gi|115249560	36	2	0	0	0	1	0
545 769	BASYS00362, 340212-339259 (CounterClockwise) Hypothetical Protein BASYS00362 Ornithine cyclodeaminase	stb_BASYS00362	37	0	0	1	0	0	0

	Identification	Number	Mass (kDa)	No. unique peptides					
				A		B-1		Tra5/5T	
				1	2	1	2	1	2
	Strains A and B-1	M = 70, s/v = 13, h = 23, R = 19, t= 8, r = 11							
	BASYS00991, 1026479-1033579 (Clockwise) Hypothetical Protein BASYS00991	sta_BASYS00991 (115249675)	269	29	18	11	9	0	0
15 399	(toxin B) toxin B	630_gi|115249675	270	0	0	3	3	0	0
88	putative subunit of oxidoreductase	630_gi|115249126	27	8	1	4	2	0	0
125	stage IV sporulation protein A	630_gi|115251680	56	8	3	0	1	0	0
137	putative 5-nitroimidazole reductase	630_gi|115250500	18	7	4	1	2	0	0
151	conserved hypothetical protein	630_gi|115250755	12	4	0	3	2	0	0
166	single-strand binding protein	630_gi|115252727	16	4	4	4	4	0	0
167	conserved hypothetical protein	630_gi|115252018	11	3	5	4	2	0	0
215	BASYS01047, 1100413-1101000 (Clockwise) HDIG Domain Protein (putative phosphohydrolase)	sta_BASYS01047 (115249720)	22	2	1	1	1	0	0
222	ABC transporter, substrate-binding protein	630_gi|115252047	35	1	3	3	3	0	0
230	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252465	18	3	2	2	2	0	0
232	putative ATPase	630_gi|115251304	32	3	0	6	3	0	0
254	anti-sigma F factor antagonist	630_gi|115249787	13	1	4	2	2	0	0
277	putative exported protein	630_gi|115250249	32	1	1	1	1	0	0
278	DNA-directed RNA polymerase omega chain	630_gi|115251644	10	1	0	4	3	0	0
279	50S ribosomal protein L22	630_gi|115249082	12	1	0	4	4	0	0
280	putative regulatory protein	630_gi|115250940	19	5	2	1	0	0	0
281	BASYS00907, 938664-938939 (Clockwise) Hypothetical Protein BASYS00907 (conserved hypothetical protein)	sta_BASYS00907 (115249602)	10	1	1	1	0	0	0
286	probable hydrolase	630_gi|115251629	24	3	1	2	2	0	0
304	putative exported protein	630_gi|115252734	21	2	1	1	1	0	0
305	putative reductase	630_gi|115249773	24	4	0	1	2	0	0
311	putative tRNA/rRNA methyltransferase	630_gi|115249058	27	1	1	0	1	0	0
312	ABC transporter, permease protein	630_gi|115249889	32	2	0	2	2	0	0
313	putative ferrous iron transport protein B	630_gi|115252330	71	3	0	2	3	0	0
314	putative phosphoribosylaminoimidazole-succinocarboxamide synthase	630_gi|115249503	25	3	0	1	2	0	0
315	ferrous iron transport protein B	630_gi|115250520	78	1	0	3	4	0	0
331	succinyl-CoA:coenzyme A transferase	630_gi|115251398	56	4	2	1	2	0	0
336	thiH, 2292548-2293924 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	stt_BASYS02217 (115251209)	52	1	1	3	1	0	0
337	putative oligopeptide transporter	630_gi|115250443	67	3	0	2	2	0	0
338	conserved hypothetical protein	630_gi|115249879	33	1	0	1	1	0	0
339	putative regulatory protein	630_gi|115250392	18	2	0	2	2	0	0
341	PTS system, IIb component	630_gi|115252071	11	2	0	1	4	0	0
342	putative beta-lactamase repressor	630_gi|115251442	15	1	0	2	1	0	0
352	probable amino acid racemase	630_gi|115251082	26	2	1	2	2	0	0
355	dipicolinate synthase, B chain	630_gi|115252025	21	3	1	1	1	0	0
356	tellurium resistance protein	630_gi|115250845	21	2	0	2	2	0	0
357	conserved hypothetical protein	630_gi|115250233	13	2	0	1	2	0	0

358	flagellar hook protein	630_gi|115249263	35	3	0	4	1	0	0
480	flagellar basal body-associated protein	630_gi|115249267	17	1	0	1	1	0	0
363	putative membrane protein	630_gi|115252518	39	3	0	1	0	0	0
371	putative nucleic acid-binding protein	630_gi|115250365	59	3	1	1	1	0	0
373	ABC transporter, ATP-binding/permease protein	630_gi|115251265	83	2	0	2	2	0	0
374	ATP synthase subunit delta	630_gi|115252531	21	2	0	1	2	0	0
375	putative membrane protein	630_gi|115249607	14	1	0	2	2	0	0
376	putative exported protein	630_gi|115249010	11	1	0	2	1	0	0
377	histidine biosynthesis bifunctional protein includes: phosphoribosyl-AMP cyclohydrolase and phosphoribosyl-ATP pyrophosphatase	630_gi|115250595	26	1	0	2	3	0	0
390	putative cardiolipin synthetase	630_gi|115252464	57	2	1	1	1	0	0
392	cobalt-precorrin-8x methylmutase	630_gi|115252490	23	1	1	2	1	0	0
393	conserved hypothetical protein	630_gi|115251576	13	1	1	1	1	0	0
394	conserved hypothetical protein	630_gi|115249575	22	2	0	1	1	0	0
395	rpmC, 425747-425962 (Clockwise) Hypothetical Protein rpmC (50S ribosomal protein L29)	sta_BASYS00420 (115249085)	8	1	2	2	0	0	0
396	trans-2-enoyl-ACP reductase	630_gi|115250213	33	1	0	1	1	0	0
404	DNA polymerase III alpha subunit	630_gi|115252456	136	1	1	2	1	0	0
405	ribose-5-phosphate isomerase 1	630_gi|115251375	16	2	1	1	1	0	0
409	ABC transporter, ATP-binding protein	630_gi|115250569	25	1	1	1	1	0	0
412	50S ribosomal protein L7/L12	630_gi|115249068	13	3	0	1	2	0	0
413	phosphoribosylaminoimidazole carboxylase catalytic subunit	630_gi|115249226	17	2	1	0	2	0	0
414	50S ribosomal protein L18	630_gi|115249093	14	1	0	1	1	0	0
415	conserved hypothetical protein	630_gi|115249462	14	1	0	2	1	0	0
417	hypothetical protein	630_gi|115252039	74	1	0	2	2	0	0
418	50S ribosomal protein L17	630_gi|115249107	13	1	0	1	1	0	0
421	aspS, 3407936-3406149 (CounterClockwise) Hypothetical Protein asps (putative aspartyl-tRNA synthetase)	sta_BASYS03216 (115251791)	67	3	1	0	1	0	0
422	cell surface protein	630_gi|115251764	72	2	0	0	1	0	0
436	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	630_gi|115249132	45	1	1	1	1	0	0
437	xanthine phosphoribosyltransferase	630_gi|115251385	21	1	1	1	1	0	0
441	glutamyl-tRNA synthetase	630_gi|115249054	57	1	1	1	1	0	0
445	conserved hypothetical protein	630_gi|115251612	20	1	1	2	0	0	0
446	conserved hypothetical membrane protein	630_gi|115250847	31	1	0	2	1	0	0
448	electron transport complex protein	630_gi|115250169	35	1	1	0	1	0	0
449	yhgF, 142962-145112 (Clockwise) Hypothetical Protein yhgF (putative RNA-binding protein)	stt_BASYS00148 (115249154)	81	3	0	0	1	0	0
450	BASYS02718, 2834398-2834129 (CounterClockwise) Hypothetical Protein BASYS02718 (conserved hypothetical protein)	stt_BASYS02718 (115251678)	10	0	2	1	0	0	0
452	V-type sodium ATP synthase subunit C	630_gi|115252014	37	1	0	4	0	0	0
456	V-type sodium ATP synthase subunit G	630_gi|115252013	12	1	1	1	1	0	0
463	50S ribosomal protein L30	630_gi|115249095	7	1	1	1	2	0	0

464	fabD, 1551856-1552806 (Clockwise) Hypothetical Protein fabD malonyl CoA-acyl carrier protein transacylase)	sta_BASYS01457 (115250214)	34	2	1	0	1	0	0
466	yaeC, 1591483-1592274 (Clockwise) Hypothetical Protein yaeC (putative D-methionine ABC transporter, substrate binding lipoprotein)	stt_BASYS01553 (11525032)	29	0	1	2	2	0	0
467	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	630_gi|115249051	17	2	1	1	0	0	0
469	ATP-dependent nuclease subunit B	630_gi|115250061	134	1	0	1	1	0	0
470	GntR-family transcriptional regulator	630_gi|115250423	24	1	0	1	1	0	0
471	putative sugar transporter, substrate-binding lipoprotein	630_gi|115251604	48	1	0	1	1	0	0
472	BASYS03031, 3208770-3207811 (CounterClockwise) Hypothetical Protein CTC (conserved hypothetical protein)	stt_BASYS03031 (115252038)	37	1	0	1	1	0	0
473	putative peptidoglycan-binding exported protein	630_gi|115252020	52	1	0	1	3	0	0
475	hypothetical protein	630_gi|115249824	14	1	0	1	1	0	0
476	rplS, 1334033-1334383 (Clockwise) 50S ribosomal protein L19	stt_BASYS01301	13	2	0	0	2	0	0
479	rpsF, 4075279-4075001 (CounterClockwise) 30S ribosomal protein S6	stt_BASYS03813	?	0	2	1	0	0	0
483	phosphopantetheine adenylyltransferase	630_gi|115251613	18	2	0	1	0	0	0
485	BASYS02045, 2108881-2108549 (CounterClockwise) Hypothetical Protein Mbar A (putative decarboxylase)	stt_BASYS02045 (115251040)	12	0	2	1	0	0	0
495	BASYS00644, 660422-660628 (Clockwise) Hypothetical Protein LJ (conserved hypothetical protein)	sta_BASYS00644 (115249289)	8	0	2	1	0	0	0
499	putative rod shape-determining protein	630_gi|115250178	32	0	2	1	2	0	0
500	putative ethanolamine/propanediol utilization protein	630_gi|115250972	18	0	1	1	2	0	0
505	50S ribosomal protein L2	630_gi|115249080	30	1	0	1	1	0	0
507	rod shape-determining protein	630_gi|115249136	36	1	0	1	1	0	0
508	probable protease	630_gi|115252340	37	1	0	1	1	0	0
510 659	putative acetyltransferase BASYS01487, 1581279-1580758 (CounterClockwise) Hypothetical Protein BASYS01487 putative acetyltransferase	630_gi|115250244 sta_BASYS01487 (115250244)	18 20	1 1	0 0	1 0	1 0	0 0	0 0
511	putative ATP phosphoribosyltransferase	630_gi|115250589	23	1	0	1	1	0	0
512	cytidylate kinase	630_gi|115250861	24	1	0	1	1	0	0
513	putative acyltransferase	630_gi|115250862	23	1	0	1	1	0	0
514	conserved hypothetical protein	630_gi|115250980	14	1	0	1	1	0	0
515	GTP-binding protein	630_gi|115251490	34	1	0	1	1	0	0
517	conserved hypothetical protein	630_gi|115250815	159	1	1	1	2	0	0
518	630_gi|115250181|emb|CAJ68002.1|-R	630_gi|115250181	?	1	0	1	1	0	0
519	anti-sigma-B factor (serine-protein kinase)	630_gi|115249013	15	1	0	1	1	0	0
522	cobyric acid synthase	630_gi|115252495	57	0	2	2	0	0	0
523	30S ribosomal protein S17	630_gi|115249086	10	1	0	1	1	0	0
529	ribosome-binding factor A	630_gi|115250346	14	1	0	1	1	0	0
530	putative proline racemase	630_gi|115252294	36	2	0	1	0	0	0
532	nitrilase (carbon-nitrogen hydrolase)	630_gi|115251895	34	2	0	1	0	0	0

540	ADP-ribose pyrophosphatase	630_gi|115250253	20	1	0	2	0	0	0
543	DNA gyrase subunit B	630_gi|115249008	71	1	0	0	2	0	0
546	probable carbohydrate hydrolase (N-terminus)	630_gi|115252061	11	0	1	1	1	0	0
548	4-hydroxyphenylacetate decarboxylase, regulatory subunit	630_gi|115249163	10	0	1	2	0	0	0
553	GntR-family transcriptional regulator	630_gi|115249663	26	1	1	1	1	0	0
554	putative xanthine/uracil permease	630_gi|115251161	48	1	0	2	1	0	0
555	cell surface protein	630_gi|115249861	34	1	1	2	0	0	0
559	PTS system, IIbc component	630_gi|115251563	56	1	0	1	1	0	0
561	putative FolD bifunctional protein includes: methylenetetrahydrofolate dehydrogenase; methenyltetrahydrofolate cyclohydrolase	630_gi|115249737	31	2	0	1	0	0	0
565	wrbA, 1228968-1229501 (Clockwise) Hypothetical Protein wrbA (putative flavodoxin))	sta_BASYS01160 (115249827)	20	2	0	0	1	0	0
568	probable peptidase	630_gi|115251664	40	2	0	1	0	0	0
573	GntR-family transcriptional regulator	630_gi|115250658	14	2	0	0	1	0	0
597	putative membrane protein	630_gi|115250318	80	0	2	1	0	0	0
598	ferric uptake regulation protein	630_gi|115250321	18	1	0	0	1	0	0
599	putative endoribonuclease	630_gi|115250431	13	1	0	0	1	0	0
600	two-component response regulator	630_gi|115250731	26	1	0	0	1	0	0
601	putative copper-transporting P-type ATPase	630_gi|115251169	89	1	0	0	1	0	0
602	putative regulatory protein	630_gi|115251736	67	1	0	0	1	0	0
603	ybaK, 1165955-1166422 (Clockwise) YbaK/Prolyl-TRNA Synthetase Associated Region	stt_BASYS01124	17	1	0	0	1	0	0
623	putative polysaccharide deacetylase	630_gi|115250355	28	1	0	0	1	0	0
629	ydfH, 1049437-1050105 (Clockwise) Hypothetical Protein ydfH (GntR-family transcriptional regulator)	stt_BASYS01029 (115249901)	26	0	1	0	1	0	0
630	putative aconitase/3-isopropylmalate dehydratase	630_gi|115249850	69	0	1	1	0	0	0
634	putative membrane protein	630_gi|115250109	42	1	0	0	1	0	0
635	branched chain amino acid transport system carrier protein	630_gi|115250294	45	1	0	0	1	0	0
636	ABC transporter, permease protein	630_gi|115250568	45	1	0	0	1	0	0
642	putative nitroreductase	630_gi|115252261	22	2	0	1	0	0	0
643	aspartokinase	630_gi|115251108	43	1	0	1	0	0	0
648	conserved hypothetical protein	630_gi|115251027	11	0	2	1	0	0	0
672	PTS system, lichenan-specific IIa component	630_gi|115251933	12	0	1	1	0	0	0
675	two-component response regulator	630_gi|115250828	28	1	0	1	0	0	0
676	conserved hypothetical protein	630_gi|115252306	21	1	0	1	0	0	0
677	undecaprenyl pyrophosphate synthetase	630_gi|115251189	28	1	0	0	1	0	0
680	putative methyl accpeting chemotaxis protein	630_gi|115249554	62	0	1	0	1	0	0

		630_gi|115250283	137	0	0	1	1	0	0
	chromosome partition protein	sta_BASYS01540							
579	smc, 1623439-1626993 (Clockwise)	115250283							
528	Chromosome partition protein smc		137	3	1	0	0	0	0
	BASYS03070, 3234507-3232624 (CounterClockwise) Hypothetical Protein BASYS03070	stb_BASYS03070 (115251842)	68	0	0	1	1	0	0
	Cell surface protein BASYS03265, 3471128-3469293 (CounterClockwise) Hypothetical Protein								
588	BASYS03265	sta_BASYS032651(							
282	Cell surface protein	115251842)	67	4	2	0	0	0	0
	BASYS03018, 3173036-3171459 (CounterClockwise) Hypothetical Protein BASYS03018	stb_BASYS03018 (115251839)	57	0	0	0	1	0	0
811	Cell surface protein								
752	cell surface protein	630_gi|115251839	57	1	0	0	0	0	0
	N-acetylmuramoyl-L-alanine amidase (cell wall hydrolase) (partial) region relative to most database matches.	630_gi|115251237	20	4	3	0	0	0	0
318	BASYS01205, 1238455-1239267 (Clockwise)	stb_BASYS01205							
477	N-Acetylmuramoyl-L-Alanine Amidase	(115251237)	30	0	0	2	2	0	0
	BASYS01676, 1758384-1763561 (Clockwise) Hypothetical Protein BASYS01676	sta_BASYS01676 (115250412)	191	1	2	0	0	0	0
	Hypothetical protein RPGR, 1516218-1521539 (Clockwise) X-	stb_BASYS01490							
570	linked retinitis pigmentosa GTPase	115250412							
790	regulator		198	0	0	1	0	0	0
	Identification	Number	Mass	No. unique peptides					
				A		B-1		Tra5/5	
				1	2	1	2	1	2
	Strains A and Tra5/5	M = 15, s/v = 2, h = 2, R = 1, t= 1, r = 0							
	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269	sta_BASYS03269 (115251852)	80	42	36	0	0	15	15
	(cell surface protein) BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898	stt_BASYS02898							
69	(Cell surface protein)	(115251852)	80	0	0	0	0	9	9
28	amino acid ABC transporter, substrate-binding protein	630_gi|115250820	30	11	9	0	0	1	0
	BASYS03267, 3473846-3471975 (CounterClockwise) Hypothetical Protein BASYS03267	sta_BASYS03267 (115251844)	66	19	11	0	0	6	9
	(cell surface protein(putative S-layer protein precursor) BASYS02896, 3045252-3043381 (CounterClockwise) Hypothetical Protein BASYS02896								
59	Cell surface protein (putative S-layer	stt_BASYS02896							
793	protein precursor)	(115251844)	66	0	0	0	0	1	0
231	gcvP, 2151957-2153414 (Clockwise)	sta_BASYS02061	54	0	1	0	0	2	3

	Hypothetical Protein gcvP (glycine cleavage system P protein)	(115250699)							
265	putative GTP-binding protein	630_gi|115251683	50	0	1	0	0	1	6
291	putative carbon-nitrogen hydrolase	630_gi|115249501	31	2	3	0	0	3	1
457	putative pyridine nucleotide-disulfide oxidoreductase	630_gi|115250843	59	1	1	0	0	2	1
461	3-oxoacyl-acyl-carrier protein reductase	630_gi|115250215	27	1	1	0	0	1	1
468	mreB, 1233049-1234059 (Clockwise) Hypothetical Protein mreB (rod shape-determining protein)	stt_BASYS01195 (115250177)	36	2	1	0	0	1	0
501	conserved hypothetical protein	630_gi|115252618	28	2	0	0	0	1	1
502	aspartate--ammonia ligase	630_gi|115251746	39	1	1	0	0	1	0
506	glycyl-tRNA synthetase alpha chain	630_gi|115251486	34	0	1	0	0	1	1
516	putative oxidoreductase subunit	630_gi|115251479	20	1	1	0	0	1	0
520	adenine phosphoribosyltransferase	630_gi|115251797	19	1	0	0	0	1	1
596	putative DNA translocase	630_gi|115250360	88	0	1	0	0	0	1
633	serine acetyltransferase	630_gi|115250636	22	0	1	0	0	0	1
641	ybgG, 3687856-3685256 (CounterClockwise) Hypothetical Protein ybgG (putative glycosyl hydrolase)	sta_BASYS03452 (115252069)	100	1	0	0	0	0	1
646	putative amino acid amidase	630_gi|115250435	30	0	1	0	0	0	1
679	ydcR, 2601886-2600678 (CounterClockwise) Hypothetical Protein ydcR (putative aminotransferase)	stt_BASYS02505 (115251476)	45	0	1	0	0	1	0
685	conserved hypothetical protein	630_gi|115249738	24	1	0	0	0	0	1
	crr, 3807137-3806646 (CounterClockwise) Hypothetical Protein crr	sta_BASYS03551 (115252157)	18	1	0	0	0	0	0
859 898	putative PTS syste, IIa component [putative PTS syste, IIa component	630_gi|115252157	18	0	0	0	0	1	0
	Identification	Number	Mass	No. unique peptides					
				A		B-1		Tra5/5	
				1	2	1	2	1	2
	Strains B-1 and Tra5/5	M = 16, s/v = 2, h = 4, R = 3, t= 2, r = 0							
188	60 chaperonin	630_gi|115249204	58	0	0	2	2	2	2
233	glycine/sarcosine/betaine reductase complex component C alpha subunit	630_gi|115251403	41	0	0	4	6	0	1
242	BASYS00004, 3075-2170 (CounterClockwise) ATPase Of	stbphage2_BASYS 00004	33	0	0	5	5	0	2
334	putative acetyltransferase	630_gi|115250223	20	0	0	3	2	1	1
368	putative sulfonate ABC transporter,solute-binding lipoprotein	630_gi|115251417	40	0	0	1	1	0	1
403	conserved hypothetical protein	630_gi|115250029	32	0	0	1	2	1	1
407	chaperone protein	630_gi|115251514	42	0	0	1	1	1	1

	cell surface protein (putative hemagglutinin/adhesin) BASYS00903, 928449-925966 (CounterClockwise) Hypothetical Protein BASYS00903	630_gi|115249532	167	0	0	0	1	1	2
410 810	Cell surface protein (putative hemagglutinin/adhesin	stb_BASYS00903 (115249532)	86	0	0	0	1	0	0
416	BASYS01206, 1239342-1239710 (Clockwise) Hypothetical Protein BASYS01206	stb_BASYS01206	14	0	0	1	1	1	0
443	putative aspartate aminotransferase	630_gi|115251434	47	0	0	2	0	1	2
444	putative oxidoreductase, NAD/FAD binding subunit	630_gi|115249186	46	0	0	1	2	1	0
447	pflD, 151794-154502 (Clockwise) Hypothetical Protein pflD 4-hydroxyphenylacetata decarboxylase, catalytic subunit	stt_BASYS00156 (115249162)	101	0	0	0	1	2	1
462	putative selenocysteine lyase	630_gi|115252735	42	0	0	1	1	1	1
498	putative zinc-binding dehydrogenase	630_gi|115251896	36	0	0	1	1	1	1
509	guanylate kinase	630_gi|115251645	23	0	0	1	1	0	1
521	conserved hypothetical protein	630_gi|115251116	38	0	0	1	1	0	1
538	porphyrin biosynthesis protein includes: uroporphyrinogen-III methyltransferase and uroporphyrinogen-III synthase	630_gi|115252480	56	0	0	0	2	0	1
556	conserved hypothetical protein	630_gi|115251115	24	0	0	1	1	1	0
557	putative extracellular solute-binding protein	630_gi|115251696	48	0	0	1	1	1	0
558	low molecular weight protein-tyrosine-phosphatase	630_gi|115252541	17	0	0	1	1	1	0
563	putative ribulose-phosphate 3-epimerase	630_gi|115251374	25	0	0	0	1	1	1
590	methionyl-tRNA formyltransferase	630_gi|115251640	34	0	0	0	1	1	0
614	Spo0B-associated GTP-binding protein	630_gi|115250196	47	0	0	1	1	0	1
632	dihydrodipicolinate reductase	630_gi|115252286	28	0	0	0	1	0	1
647	GTP-binding elongation factor	630_gi|115251521	67	0	0	1	0	1	0
674	amino acid ABC transporter, substrate-binding protein	630_gi|115249767	30	0	0	0	1	0	1
	putative prephenate dehydrogenase	630_gi|115250885	31	0	0	1	0	0	0
813 631	BASYS01924, 1985976-1986815 (Clockwise) Hypothetical Protein BASYS01924 Putative prephenate dehydrogenase	stt_BASYS01924 (115250885)	31	0	0	0	0	1	1
	Identification	Number	Mass	No. unique peptides					
				A		B-1		Tra5/5	
				1	2	1	2	1	2
	Strain Tra5/5 only	M = 33, s/v = 1, h = 6, R = 3, t= 2, r = 0							
295	seryl-tRNA synthetase	630_gi|115249017	49	0	0	0	0	1	1
316	fdhF, 2319506-2317329 (CounterClockwise) Hypothetical Protein fdhF (putative formate dehydrogenase)	stt_BASYS02242 (115251232)	83	0	0	0	0	2	2
360	yiaJ, 3855532-3856284 (Clockwise) Hypothetical Protein yiaJ	stt_BASYS03597	?	0	0	0	0	3	4
362	putative molybdenum-binding subunit of oxidoreductase	630_gi|115251153	82	0	0	0	0	2	4
378	putative electron transfer protein	630_gi|115252303	47	0	0	0	0	2	3
380	asparaginyl-tRNA synthetase	630_gi|115251299	54	0	0	0	0	2	4

451	cysS, 393517-394923 (Clockwise) Hypothetical Protein cysS cysteinyl-tRNA synthetase	sta_BASYS00384 (115249055)	54	0	0	0	0	2	2
481	putative chorismate biosynthesis-related protein includes: phospho-2-dehydro-3-deoxyheptonate aldolase and chorismate mutase	630_gi|115250878	37	0	0	0	0	2	1
527	putative FAD-binding subunit of oxidoreductase	630_gi|115251155	30	0	0	0	0	1	2
539	dihydroorotate dehydrogenase, catalytic subunit	630_gi|115249196	32	0	0	0	0	2	1
542	putative aminopeptidase	630_gi|115250096	51	0	0	0	0	1	1
547	threonine synthase	630_gi|115251172	56	0	0	0	0	1	2
572	putative radical SAM family protein	630_gi|115251801	53	0	0	0	0	1	1
591	phosphoglucomutase/phosphomannomutase mutase	630_gi|115249128	51	0	0	0	0	1	1
592	conserved hypothetical protein	630_gi|115250761	39	0	0	0	0	1	1
593	N-carbamoyl-L-amino acid hydrolase	630_gi|115251081	44	0	0	0	0	1	1
594	putative thymidylate kinase	630_gi|115252611	27	0	0	0	0	1	1
595	pyrroline-5-carboxylate reductase	630_gi|115252337	28	0	0	0	0	1	0
606	chorismate synthase (5-enolpyruvylshikimate-3-phosphate phospholyase)	630_gi|115250881	39	0	0	0	0	1	1
658	putative ATP/GTP-binding protein	630_gi|115249805	48	0	0	0	0	0	2
690	P-protein includes: chorismate mutase and prephenate dehydratase	630_gi|115250882	46	0	0	0	0	1	1
691	UDP-N-acetylmuramate--L-alanine ligase	630_gi|115252579	50	0	0	0	0	0	1
707	dihydrodipicolinate reductase	630_gi|115252283	27	0	0	0	0	1	0
728	radical SAM-superfamily protein	630_gi|115250362	51	0	0	0	0	0	1
729	glutamine synthetase	630_gi|115250379	72	0	0	0	0	0	1
730	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alan ine ligase	630_gi|115251706	51	0	0	0	0	0	1
731	putative GTP pyrophosphokinase	630_gi|115251796	85	0	0	0	0	0	1
732	galU, 3885914-3884934 (CounterClockwise) Hypothetical Protein galU	stt_BASYS03633	?	0	0	0	0	0	1
760	oppF, 1277546-1278490 (Clockwise) Hypothetical Protein oppF ABC-type oligopeptide transport system, ARPase component	sta_BASYS01208	36	0	0	0	0	1	1
791	putative 2-nitropropane dioxygenase	630_gi|115251174	38	0	0	0	0	1	0
792	potG, 466136-466891 (Clockwise) Hypothetical Protein potG ABC transporter, ATP-binding protein	stb_BASYS00509 (115249891)	28	0	0	0	0	1	0
795	putative DNA mismatch repair protein	630_gi|115249725	88	0	0	0	0	0	1
822	DNA repair protein	630_gi|115249030	50	0	0	0	0	0	1
823	putative hydantoinase	630_gi|115250762	56	0	0	0	0	0	1
824	PhoH-like protein	630_gi|115251494	38	0	0	0	0	0	1
825	cinserved hypothetical protein	630_gi|115251677	47	0	0	0	0	0	1
826	probable peptidase	630_gi|115251712	48	0	0	0	0	0	1

829	yibO, 3925985-3924453 (CounterClockwise) Hypothetical Protein yibO 2,3-bisphosphoglycerate-independent phosphoglycerate mutase	sta_BASYS03659 (115252228)	56	0	0	0	0	0	1
830	purF, 231027-232394 (Clockwise) Hypothetical Protein purF amidophosphoribosyltransferase	stt_BASYS00227 (115249228)	51	0	0	0	0	0	1
866	hydrogenase	630_gi|115252370	51	0	0	0	0	0	1
873	conserved hypothetical protein	630_gi|115250624	11	0	0	0	0	1	0
877	hsdM, 3719141-3716964 (CounterClockwise) Hypothetical Protein hsdM type I restriction-modification system specificity subunit	stt_BASYS03486	83	0	0	0	0	1	0
882	cyclopropane-fatty-acyl-phospholipid synthase	630_gi|115249187	46	0	0	0	0	0	1
895	putative deoxyribonuclease	630_gi|115252600	29	0	0	0	0	1	0
925	cell surface protein	630_gi|115251246	51	0	0	0	0	0	1
	Identification	Number	Mass	No. unique peptides					
				A		B-1		Tra5/5	
				1	2	1	2	1	2
	Strain B-1 only	M = 68, s/v = 12, h = 56, R = 17, t= 5, r = 3							
33	BASYS03072, 3239516-3237684 (CounterClockwise) Hypothetical Protein BASYS03072 Cell surface protein (S-layer precursor protein)	stb_BASYS03072 (115251846)	65	0	0	24	20	0	0
168	putative carbon-nitrogen hydrolase	630_gi|115251789	31	0	0	7	6	0	0
255	transcription antiterminator	630_gi|115252086	32	0	0	4	5	0	0
317	BASYS00003, 1394-2224 (Clockwise) Hypothetical Protein BASYS00003	Stbphage1_BASYS 00003	33	0	0	3	4	0	0
343	conserved hypothetical protein	630_gi|115252214	30	0	0	3	2	0	0
344	ybhF, 1494368-1495099 (Clockwise) Hypothetical Protein ybhF putative lantibiotic ABC transporter,ATP-binding protein	stb_BASYS01462 (115250385)	27	0	0	3	4	0	0
361	sopA, 5108-4290 (CounterClockwise) Hypothetical Protein sopA	Stbphage1_BASYS 00008	32	0	0	2	4	0	0
379	ybaB, 338454-338834 (Clockwise) Hypothetical Protein ybaB conserved hypothetical protein	sta_BASYS00340 (115249020)	13	0	0	3	3	0	0
397	putative oxidoreductase, electron transfer subunit	630_gi|115249185	18	0	0	2	1	0	0
398	putative L-threonine dehydrogenase	630_gi|115250838	36	0	0	2	2	0	0
858	galE, 2294153-2293197 (CounterClockwise) Hypothetical Protein gale putative L-threonine dehydrogenase	sta_BASYS02203 (115250838)	36	1	0	0	0	0	0
420	BASYS02421, 2523005-2522130 (CounterClockwise) Hypothetical Protein BASYS02421 Hypothetical protein	sta_BASYS02421 (115251013)	34	0	0	2	2	0	0

423	BASYS00597, 615300-615635 (Clockwise) Hypothetical Protein BASYS00597 Hypothetical protein	sta_BASYS00597 (115249246)	13	0	0	3	3	0	0
425	BASYS00012, 9275-9622 (Clockwise) Hypothetical Protein BASYS00012	Stbphage1_BASYS 00012	14	0	0	2	3	0	0
426	putative phosphosugar isomerase	630_gi|115252083	21	0	0	2	2	0	0
429	dihydropteroate synthase	630_gi|115250491	30	0	0	3	1	0	0
430	conserved hypothetical protein	630_gi|115249635	9	0	0	1	0	0	0
453	PTS system, IIabc component	630_gi|115252143	52	0	0	1	1	0	0
474	putative 3-mercaptopyruvate sulfurtransferase	630_gi|115250576	32	0	0	2	2	0	0
482	BASYS01010, 1054756-1055364 (Clockwise) Hypothetical Protein BASYS01010 Conserved hypothetical protein	sta_BASYS01010 (115249688)	23	0	0	1	3	0	0
484	BASYS00060, 38456-38932 (Clockwise) Hypothetical Protein BASYS00060 Hypothetical protein	stbphage2_BASYS 00060	17	0	0	1	2	0	0
486	thioredoxin	630_gi|115251408	12	0	0	1	2	0	0
487	conserved hypothetical protein	630_gi|115249636	8	0	0	2	1	0	0
494	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	630_gi|115252499	37	0	0	1	0	0	0
503	conserved hypothetical protein	630_gi|115251009	30	0	0	1	1	0	1
504	peptide chain release factor 1	630_gi|115252544	40	0	0	1	0	0	0
524	BASYS02039, 2122887-2123531 (Clockwise) Hypothetical Protein BASYS02039 Tellurium resistance protein	sta_BASYS02039 (115250677)	24	0	0	2	2	0	0
525	rRNA adenine N-6-methyltransferase (erythromycin resistance protein)	630_gi|115251058	29	0	0	2	2	0	0
531	putative membrane protein	630_gi|115249794	29	0	0	1	2	0	0
533	conserved hypothetical protein	630_gi|115252657	22	0	0	1	2	0	0
534	BASYS00477, 430146-431189 (Clockwise) Hypothetical Protein BASYS00477 No c. difficile match by BLAST	stb_BASYS00477	41	0	0	1	2	0	0
535	thioredoxin reductase	630_gi|115250734	33	0	0	1	2	0	0
536	putative ATP-binding protein	630_gi|115250457	25	0	0	1	0	0	0
541	BASYS03056, 3218968-3215420 (CounterClockwise) Hypothetical Protein BASYS03056	stb_BASYS03056	139	0	0	1	2	0	0
544	PTS system, IIa component	630_gi|115251565	17	0	0	1	1	0	0
562	cobalt-precorrin-6a reductase; precorrin-6x reductase	630_gi|115252483	28	0	0	2	2	0	0
564	signal recognition particle protein	630_gi|115250286	50	0	0	1	1	0	0
566	50S ribosomal protein L23	630_gi|115249079	11	0	0	1	2	0	0
567	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	630_gi|115250983	32	0	0	1	2	0	0
571	putative flagellar motor switch protein	630_gi|115249279	33	0	0	1	1	0	0
702	fliN, 608925-609239 (Clockwise) Hypothetical Protein fliN flagellar motor switch protein	sta_BASYS00587 (115249236)	12	0	0	1	0	0	0
710	putative flagellar motor switch protein	630_gi|115249280	14	0	0	0	1	0	0
575	conserved hypothetical protein	630_gi|115250343	11	0	0	1	1	0	0
576	probable esterase	630_gi|115251258	30	0	0	1	1	0	0
577	conserved hypothetical protein	630_gi|115249911	14	0	0	1	1	0	0
578	50S ribosomal protein L27	630_gi|115250195	10	0	0	1	1	0	0

580	pyrroline-5-carboxylate reductase	630_gi|115250536	29	0	0	1	1	0	0
581	glycine/sarcosine/betaine reductase complex component A	630_gi|115251406	17	0	0	1	1	0	0
582	putative transcription antiterminator	630_gi|115251719	33	0	0	1	1	0	0
583	cell surface protein	630_gi|115251849	67	0	0	1	1	0	0
584	PTS system, maltose and glucose-specific IIbc component	630_gi|115252085	57	0	0	1	1	0	0
585	putative stage 0 sporulation protein	630_gi|115252738	30	0	0	1	1	0	0
586	adhE, 2136326-2137795 (Clockwise) Hypothetical Protein adhE putative ethanolamine/propanediol utilisation aldehyde-alcohol dehydrogenase	stb_BASYS02079 (115250964)	53	0	0	1	1	0	0
587	BASYS03023, 3180249-3179710 (CounterClockwise) Hypothetical Protein Lmo Conjugative transposon conserved hypothetical protein	stb_BASYS03023 (115249366)	21	0	0	1	1	0	0
589	BASYS03016, 3189040-3188357 (CounterClockwise) Hypothetical Protein BF Conserved hypothetical protein	stt_BASYS03016 (115252021)	26	0	0	1	1	0	0
	BASYS00474, 429093-429458 (Clockwise) Hypothetical Protein BASYS00474	stb_BASYS00474 (115251071)	14	0	0	1	1	0	0
604 381	Putative phage regulatory protein (XRE family transcriptional regulator) BASYS00015, 10793-10440 (CounterClockwise) Hypothetical Protein BASYS00015 Transcriptional regulator, xre family protein	Stbphage1_BASYS 00015	14	0	0	2	2	0	0
611	transcription elongation factor	630_gi|115252616	18	0	0	2	0	0	0
612	30S ribosomal protein S16	630_gi|115250287	10	0	0	2	0	0	0
613	flavodoxin	630_gi|115251047	16	0	0	2	0	0	0
619	rffG, 606774-607757 (Clockwise) Hypothetical Protein rffG (dTDP-glucose 4,6-dehydratase)	sta_BASYS00584	38	0	0	1	1	0	0
620	homoserine kinase	630_gi|115251173	33	0	0	1	1	0	0
624	putative small-molecule-binding protein	630_gi|115249015	20	0	0	1	1	0	0
625	ATP-dependent nuclease subunit A	630_gi|115250062	147	0	0	1	1	0	0
626	putative tRNA modification GTPase	630_gi|115252740	51	0	0	1	1	0	0
627	BASYS03349, 3561478-3560975 (CounterClockwise) Hypothetical Protein BASYS03349 No match by BLAST	stb_BASYS03349	19	0	0	1	1	0	0
644	cysG, 9043-8270 (CounterClockwise) Hypothetical Protein cysG precorrin-4 C(11)-methyltransferase	sta_BASYS00009 (115252486)	29	0	0	1	1	0	0
645	BASYS00154, 142715-143050 (Clockwise) Hypothetical Protein BASYS00154 Phage protein	stb_BASYS00154 (115249958)	13	0	0	1	1	0	0
649	PTS system, IIa component	630_gi|115252072	17	0	0	1	1	0	0
650	stage V sporulation protein G	630_gi|115252577	10	0	0	2	0	0	0
651	hypothetical protein	630_gi|115252299	11	0	0	1	1	0	0
653	hypothetical protein	630_gi|115251805	26	0	0	0	1	0	0
654	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	630_gi|115250982	35	0	0	1	0	0	0

660	putative exported protein	630_gi|115251418	17	0	0	0	1	0	0
661	conserved hypothetical protein	630_gi|115251421	12	0	0	1	0	0	0
662	BASYS00184, 156070-156327 (Clockwise) Hypothetical Protein BASYS00184 No match to BLAST	stb_BASYS00184	10	0	0	1	0	0	0
663	BASYS01737, 1767940-1768269 (Clockwise) Hypothetical Protein BASYS01737 No match to BLAST	stb_BASYS01737	13	0	0	1	0	0	0
667	conserved hypothetical protein	630_gi|115249350	16	0	0	1	1	0	0
668	conserved hypothetical protein	630_gi|115249339	28	0	0	1	1	0	0
669	putative chorismate biosynthesis-related protein includes: phospho-2-dehydro-3-deoxyheptonate aldolase and chorismate mutase	630_gi|115250494	29	0	0	1	1	0	0
670	cysteine synthase A	630_gi|115250707	35	0	0	1	1	0	0
671	BASYS03069, 3232315-3231857 (CounterClockwise) Hypothetical Protein BASYS03069 No match by BLAST	stb_BASYS03069	17	0	0	1	1	0	0
682	tellurium resistance protein	630_gi|115250846	21	0	0	1	0	0	0
692	putative serine hydroxymethyltransferase	630_gi|115251778	46	0	0	1	0	0	0
693	conserved hypothetical protein	630_gi|115250202	47	0	0	1	1	0	0
696	putative phage positive regulator of late transcription	630_gi|115249641	11	0	0	1	0	0	0
697	conserved hypothetical protein	630_gi|115250817	14	0	0	1	0	0	0
698	PTS system, glucose-specific IIa component	630_gi|115252082	17	0	0	1	0	0	0
699	putative 5-formyltetrahydrofolate cyclo-ligase	630_gi|115252516	22	0	0	1	0	0	0
700	ATP synthase epsilon chain (partial)	630_gi|115252527	9	0	0	1	0	0	0
703	BASYS00188, 157642-158154 (Clockwise) Hypothetical Protein BASYS00188 Putative phage repressor	stb_BASYS00188	20	0	0	1	0	0	0
704	ydhO, 1281410-1282705 (Clockwise) Hypothetical Protein ydhO putative cell wall hydrolase	stb_BASYS01250 (115250166)	45	0	0	1	0	0	0
705 891	rfbC, 3226199-3225627 (CounterClockwise) Hypothetical Protein rfbC (dTDP-4-dehydrorhamnose 3,5-epimerase) rfbC, 606179-606736 (Clockwise) Hypothetical Protein rfbC dTDP-4-dehydrorhamnose 3,5-epimerase	stb_BASYS03063 sta_BASYS00583	22 22	0 0	0 0	1 1	0 0	0 0	0 0
706	BASYS03448, 3677647-3677129 (CounterClockwise) Hypothetical Protein BASYS03448 Putative nitroreductase	stt_BASYS03448 (115252412)	20	0	0	1	0	0	0
708	DNA polymerase III subunit gamma/tau	630_gi|115249019	63	0	0	0	1	0	0
709	dihydroorotate dehydrogenase electron transfer subunit	630_gi|115249195	26	0	0	0	1	0	0
711	putative nucleotide phosphodiesterase	630_gi|115249706	72	0	0	0	1	0	0
712	conserved hypothetical protein	630_gi|115250366	70	0	0	0	1	0	0
713	putative dinitrogenase iron-molybdenum cofactor	630_gi|115250736	13	0	0	0	1	0	0
714	ABC transporter, ATP-binding protein	630_gi|115251007	24	0	0	0	1	0	0
715	putative aromatic compounds hydrolase	630_gi|115251234	29	0	0	0	1	0	0
717	probable cobalt-precorrin-6y C(5)-methyltransferase	630_gi|115252488	23	0	0	0	1	0	0

718	carbamoyl-phosphate synthase,pyrimidine-specific, small chain	630_gi|115252653	39	0	0	0	1	0	0
719	sporulation initiation inhibitor	630_gi|115252737	28	0	0	0	1	0	0
720	BASYS00599, 616624-618843 (Clockwise) Glycosyl Transferase Group 2 Family Protein	sta_BASYS00599 (115249248)	89	0	0	0	1	0	0
721	BASYS00167, 150151-149381 (CounterClockwise) Hypothetical Protein BASYS00167 hypothetical	stb_BASYS00167	30	0	0	0	1	0	0
722	BASYS01138, 1196854-1196450 (CounterClockwise) Transcriptional Regulator Cro/CI Family	stb_BASYS01138	16	0	0	0	1	0	0
761	MarR-family transcriptional regulator	630_gi|115250333	17	0	0	0	1	0	0
766	rpoN, 3932428-3931073 (CounterClockwise) Hypothetical Protein rpoN RNA polymerase sigme-54 factor	sta_BASYS03664 (115252233)	52	0	0	1	0	0	0
775	conserved hypothetical protein	630_gi|115250436	18	0	0	0	1	0	0
776	hypothetical phage protein	630_gi|115249959	9	0	0	1	0	0	0
777	putative phage major capsid protein	630_gi|115249966	34	0	0	1	0	0	0
778	ABC transporter, substrate-binding lipoprotein	630_gi|115250020	36	0	0	1	0	0	0
779	3-oxoacyl-acyl-carrier-protein synthase III	630_gi|115250212	35	0	0	1	0	0	0
780	hypothetical protein	630_gi|115250621	19	0	0	1	0	0	0
781	conserved hypothetical protein	630_gi|115251130	42	0	0	1	0	0	0
782	conserved hypothetical protein	630_gi|115251487	22	0	0	1	0	0	0
783	putative exported protein	630_gi|115251537	39	0	0	1	0	0	0
784	coenzyme A biosynthesis bifunctional protein	630_gi|115251643	44	0	0	1	0	0	0
785	conserved hypothetical protein	630_gi|115252035	21	0	0	1	0	0	0
786	putative preprotein translocase	630_gi|115252635	19	0	0	1	0	0	0
787	conserved hypothetical protein	630_gi|115252672	33	0	0	1	0	0	0
788	gidB, 310295-309567 (CounterClockwise) Hypothetical Protein gidB methyl transferase	sta_BASYS00316	27	0	0	1	0	0	0
789	fliF, 335693-334143 (CounterClockwise) Hypothetical Protein fliF flagellar M-ring protein	stb_BASYS00356 (115249256)	56	0	0	1	0	0	0
794	trehalose-6-phosphate hydrolase	630_gi|115252145	66	0	0	1	0	0	0
796	chemotaxis protein	630_gi|115249265	30	0	0	0	1	0	0
797	conserved hypothetical protein	630_gi|115249689	27	0	0	0	1	0	0
798	hypothetical protein	630_gi|115249753	14	0	0	0	1	0	0
799	putative ribonucleotide-diphosphate reductase	630_gi|115250295	83	0	0	0	1	0	0
800	630_gi|115250446|emb|CAJ68269.1|-R	630_gi|115250446	?	0	0	0	1	0	0
801	molybdenum cofactor biosynthesis protein	630_gi|115250758	17	0	0	0	1	0	0
802	DNA mismatch repair protein	630_gi|115251022	75	0	0	0	1	0	0
804	DeoR-family transcriptional regulator	630_gi|115251165	29	0	0	0	1	0	0
805	putative phosphomannomutase/phosphoglycerate mutase	630_gi|115251833	64	0	0	0	1	0	0
806	putative DNA repair protein (nucleotide pyrophosphatase)	630_gi|115252458	16	0	0	0	1	0	0
807	excinuclease ABC subunit B	630_gi|115252472	76	0	0	0	1	0	0
808	nusA, 1689456-1690610 (Clockwise) Hypothetical Protein nusA	sta_BASYS01606 (115250342)	43	0	0	0	1	0	0

	transcription elongation protein								
809	BASYS00420, 401655-400888 (CounterClockwise) Hypothetical Protein BASYS00420 hypothetical	stb_BASYS00420	28	0	0	0	1	0	0
812	deaD, 915268-916881 (Clockwise) Hypothetical Protein dead putative ATP-dependent RNA helicase	stt_BASYS00904 (115249778)	61	0	0	0	1	0	0
814	putative ferrous iron transport protein A	630_gi|115250789	8	0	0	1	0	0	0
864	prolyl-tRNA synthetase	630_gi|115249052	64	0	0	0	1	0	0
865	transcriptional antiterminator	630_gi|115251564	32	0	0	0	1	0	0
867	putative DNA polymerase III, delta' subunit	630_gi|115252610	36	0	0	1	0	0	0
868	yjeE, 148838-149293 (Clockwise) Hypothetical Protein yjeE putative ATP/GTP hydrolase	stt_BASYS00152 (115249158)	17	0	0	0	1	0	0
874	ABC transporter, ATP-binding protein	630_gi|115250659	27	0	0	0	1	0	0
876	BASYS01324, 1412801-1414714 (Clockwise) Hypothetical Protein BASYS01324 hypothetical	sta_BASYS01324 (115250058)	75	0	0	0	1	0	0
878	phosphoglycerate kinase	630_gi|115252230	43	0	0	0	1	0	0
883	conserved hypothetical protein	630_gi|115252219	19	0	0	1	0	0	0
886	zntA, 695882-698269 (Clockwise) Hypothetical Protein zntA putative heavy-metal-transporting ATPase	sta_BASYS00689 (115249322)	86	0	0	1	0	0	0
888	chromosomal replication initiator protein	630_gi|115249004	50	0	0	1	0	0	0
890	ompR, 2456141-2455449 (CounterClockwise) Hypothetical Protein ompR two-component system response regulator	stt_BASYS02359 (115251341)	27	0	0	0	1	0	0
893	conserved hypothetical protein	630_gi|115251465	46	0	0	1	0	0	0
896	putative aminotransferase	630_gi|115251747	46	0	0	1	0	0	0
901	BASYS03068, 3231844-3230237 (CounterClockwise) Hypothetical Protein BASYS03068 No match by BLAST	stb_BASYS03068	61	0	0	1	0	0	0
902	putative flagellar biosynthesis protein	630_gi|115249238	16	0	0	1	0	0	0
904	putative helicase	630_gi|115250059	87	0	0	0	1	0	0
906	putative lipoprotein	630_gi|115250384	21	0	0	0	1	0	0
908	630_gi|115249949|emb|CAJ67769.1|-R	630_gi|115249949	?	0	0	0	1	0	0
909	putative phage-related protein	630_gi|115250413	5	0	0	1	0	0	0
910	conserved hypothetical protein	630_gi|115249410	13	0	0	1	0	0	0
911	BASYS00592, 612382-612774 (Clockwise) Hypothetical Protein BASYS00592 Hypothetical protein (Flagellar assembly factor FliW)	sta_BASYS00592 115249241	15	0	0	1	0	0	0
915	recT, 153449-152616 (CounterClockwise) Hypothetical Protein recT recombination and repair protein RecT [*Clostridium difficile* QCD-63q42]	stb_BASYS00175	31	0	0	0	1	0	0
921	putative esterase	630_gi|115250430	66	0	0	1	0	0	0
	Identification	Number	Mass (kDa)	No. unique peptides					
				A		B-1		Tra5/5	
				1	2	1	2	1	2
	Strain A only	M = 66, s/v = 14, h = 30, R = 11, t= 10, r = 1							
293	putative bifunctional protein:	630_gi|115250474	81	6	1	0	0	0	0

	peroxiredoxin/chitinase								
294	putative membrane protein	630_gi|115250974	52	8	1	0	0	0	0
306	probable proton-dependent oligopeptide transporter	630_gi|115251316	50	3	2	0	0	0	0
307	putative cell wall hydrolase	630_gi|115252630	59	5	1	0	0	0	0
319	single-stranded DNA binding protein	630_gi|115252292	25	4	1	0	0	0	0
345	BASYS01831, 1920387-1919473 (CounterClockwise) Hypothetical Protein CAC Hypothetical protein	sta_BASYS01831 (115250552)	35	4	0	0	0	0	0
382	BASYS00914, 944080-949320 (Clockwise) Hypothetical Protein Pc Hypothetical protein	sta_BASYS00914 (115249605)	199	5	1	0	0	0	0
424	BASYS03447, 3680327-3679458 (CounterClockwise) Hypothetical Protein CPE Hypothetical protein	sta_BASYS03447 (115252065)	33	2	2	0	0	0	0
428	BASYS01825, 1915638-1914313 (CounterClockwise) Hypothetical Protein BASYS01825 putative drug/sodium antiporters	sta_BASYS01825 (115250546)	49	3	2	0	0	0	0
455	glycine betaine/carnitine/choline ABC transporter, substrate-binding protein	630_gi|115249915	57	3	0	0	0	0	0
488	putative pyruvate formate-lyase 3 activating enzyme	630_gi|115252339	34	2	1	0	0	0	0
489	conserved hypothetical protein	630_gi|115252583	51	2	1	0	0	0	0
490	putative oxidoreductase	630_gi|115249453	39	2	1	0	0	0	0
491	putative esterase/halogenase	630_gi|115251917	30	2	1	0	0	0	0
492	BASYS03271, 3480933-3479344 (CounterClockwise) Hypothetical Protein BASYS03271 Surface surface protein	sta_BASYS03271 (115251847)	58	1	3	0	0	0	0
493	BASYS00861, 893869-892595 (CounterClockwise) Spore Cortex-Lytic Pre-Pro-Form	sta_BASYS00861 115249567	47	1	2	0	0	0	0
496	stage V sporulation protein T	630_gi|115252560	20	4	0	0	0	0	0
497	putative protein translocase subunit	630_gi|115251854	11	2	0	0	0	0	0
537	BASYS01931, 1999997-1998948 (CounterClockwise) Hypothetical Protein BB ABC transporter substrate-binding protein [*Clostridium difficile* QCD-63q42]	stt_BASYS01931	38	2	1	0	0	0	0
549	glnS, 2609350-2611014 (Clockwise) Glutaminyl-tRNA synthetase	sta_BASYS02511 115251113	64	3	0	0	0	0	0
550	rubrerythrin	630_gi|115251898	22	3	0	0	0	0	0
551	conserved hypothetical protein	630_gi|115252644	29	2	0	0	0	0	0
552	phosphoribosylaminoimidazole-succinocarboxamide synthase	630_gi|115249227	27	1	0	0	0	0	0
569	D-ornithine aminomutase E component	630_gi|115249457	83	2	1	0	0	0	0
608	putative sugar-phosphate kinase	630_gi|115249217	48	1	1	0	0	0	0
609	ABC transporter, ATP-binding protein	630_gi|115251510	42	2	0	0	0	0	0
610	thioredoxin	630_gi|115250733	12	0	1	0	0	0	0
618	putative manganese-containing catalase	630_gi|115250607	25	2	1	0	0	0	0
638	putative stage II sporulation protein P	630_gi|115251523	38	1	1	0	0	0	0

639	putative aromatic amino acid aminotransferase	630_gi|115251881	45	1	1	0	0	0	0
640	BASYS03656, 3921940-3921434 (CounterClockwise) Acetyltransferase GNAT Family	sta_BASYS03656	19	1	1	0	0	0	0
652	inosine-uridine preferring nucleoside hydrolase	630_gi|115250725	36	1	0	0	0	0	0
655	conserved hypothetical protein	630_gi|115251175	20	2	0	0	0	0	0
657	deoxyuridine 5'-triphosphate nucleotidohydrolase	630_gi|115251455	16	2	0	0	0	0	0
664	putative fumarate hydratase, subunit B	630_gi|115250025	20	1	0	0	0	0	0
665	putative membrane protein	630_gi|115250482	32	1	0	0	0	0	0
666	putative membrane protein	630_gi|115252127	20	1	0	0	0	0	0
678	putative ATP/GTP-binding protein	630_gi|115252354	41	1	1	0	0	0	0
683	putative penicillin-binding protein	630_gi|115250180	111	1	1	0	0	0	0
687	PTS system, IIb component	630_gi|115250107	10	0	1	0	0	0	0
688	low-specificity L-threonine aldolase	630_gi|115251648	38	2	0	0	0	0	0
689	conserved hypothetical protein	630_gi|115252296	17	1	0	0	0	0	0
694	sodium:dicarboxylate symporter family protein	630_gi|115251595	50	1	0	0	0	0	0
695	putative lipoate-protein ligase	630_gi|115250695	38	1	0	0	0	0	0
723	putative lipoprotein	630_gi|115250265	32	0	1	0	0	0	0
724	IclR-family transcriptional regulator	630_gi|115251483	29	0	1	0	0	0	0
725	putative dihydroorotate dehydrogenase, catalytic subunit	630_gi|115252236	39	0	1	0	0	0	0
726	BASYS02324, 2426262-2427068 (Clockwise) Hypothetical Protein EF Hypothetical protein	sta_BASYS02324 115250916	31	0	1	0	0	0	0
733	E3 component of acetoin dehydrogenase enzyme system (dihydrolipoyl dehydrogenase)	630_gi|115249042	62	1	0	0	0	0	0
734	flagellar export protein	630_gi|115249272	76	1	0	0	0	0	0
735	PTS system, IIb component	630_gi|115249506	17	1	0	0	0	0	0
736	putative transcriptional regulator	630_gi|115249644	25	1	0	0	0	0	0
737	conserved hypothetical protein	630_gi|115250341	18	1	0	0	0	0	0
738	putative sigma-54-dependent transcriptional regulator (partial)	630_gi|115250478	22	1	0	0	0	0	0
739	putative deoxyribose-phosphate aldolase	630_gi|115250542	24	1	0	0	0	0	0
740	putative ferrous ion transport protein	630_gi|115250558	78	1	0	0	0	0	0
741	conserved hypothetical protein	630_gi|115250804	33	1	0	0	0	0	0
742	two-component response regulator	630_gi|115250957	22	1	0	0	0	0	0
743	conserved hypothetical protein	630_gi|115251114	34	1	0	0	0	0	0
744	putative membrane protein	630_gi|115251282	30	1	0	0	0	0	0
745	putative germination-specific protease	630_gi|115251301	61	1	0	0	0	0	0
746	conserved hypothetical protein	630_gi|115251313	13	1	0	0	0	0	0
747	putative aliphatic sulfonate ABC transporter,ATP-binding protein	630_gi|115251414	30	1	0	0	0	0	0
748	conserved hypothetical protein	630_gi|115251427	11	1	0	0	0	0	0
749	histidine triad nucleotide-binding protein	630_gi|115251501	13	1	0	0	0	0	0
750	30S ribosomal protein S20	630_gi|115251527	10	1	0	0	0	0	0
751	putative phosphoribosyltransferase	630_gi|115251742	20	1	0	0	0	0	0
753	dipicolinate synthase, A chain	630_gi|115252026	32	1	0	0	0	0	0
754	delta-aminolevulinic acid dehydratase	630_gi|115252479	36	1	0	0	0	0	0
755	putative RNA/single-stranded DNA exonuclease	630_gi|115252723	76	1	0	0	0	0	0
756	thyA, 395829-396659 (Clockwise)	sta_BASYS00387	32	1	0	0	0	0	0

	Thymidylate synthase								
757	ligA, 3927560-3925527 (CounterClockwise) Hypothetical Protein ligA DNA ligase	stb_BASYS03670 (115252365)	77	2	0	0	0	0	0
758	ycjC, 1769550-1768987 (CounterClockwise) Hypothetical Protein ycjC	stt_BASYS01717 115250685	21	1	0	0	0	0	0
759	kdgR, 3236380-3237156 (Clockwise) Hypothetical Protein kdgR	stt_BASYS03056 115249418	29	1	0	0	0	0	0
763	putative exported protein	630_gi|115252517	16	1	0	0	0	0	0
767	yecS, 1915939-1916550 (Clockwise) Hypothetical Protein yecS amino acid ABC transporter, permease protein	stt_BASYS01858 (115250821)	23	2	0	0	0	0	0
770	stage V sporulation protein S	630_gi|115250981	9	0	1	0	0	0	0
771	BASYS00233, 228947-229429 (Clockwise) Glyoxalase	sta_BASYS00233	19	0	1	0	0	0	0
832	putative glyoxalase	630_gi|115249166	18	1	0	0	0	0	0
773	PTS system, mannitol-specific IIbc component	630_gi|115251389	50	0	1	0	0	0	0
815	bifunctional purine biosynthesis protein includes: phosphoribosylaminoimidazolecarboxamide formyltransferase and IMP cyclohydrolase	630_gi|115249231	57	0	1	0	0	0	0
816	methylglyoxal synthase	630_gi|115250185	15	0	1	0	0	0	0
817	putative muramoyltetrapeptide carboxypeptidase	630_gi|115250438	33	0	1	0	0	0	0
818	cell surface protein	630_gi|115251848	59	0	1	0	0	0	0
819	conserved hypothetical protein	630_gi|115252555	10	0	1	0	0	0	0
820	manX, 3750702-3750208 (CounterClockwise) Hypothetical Protein manX PTS system, IIB component	sta_BASYS03509 (11525212)	18	0	1	0	0	0	0
821	BASYS03604, 3866978-3865935 (CounterClockwise) Hypothetical Protein BASYS03604 Hypothetical protein	sta_BASYS03604	39	0	1	0	0	0	0
831	putative exported protein	630_gi|115250383	41	1	0	0	0	0	0
833	putative spore-coat protein	630_gi|115249613	21	1	0	0	0	0	0
834	conserved hypothetical protein	630_gi|115250402	17	1	0	0	0	0	0
835	putative signaling protein	630_gi|115250461	66	1	0	0	0	0	0
836	putative molybdenum cofactor biosynthesis protein	630_gi|115250756	18	1	0	0	0	0	0
837	putative glutamyl-aminopeptidase	630_gi|115251205	39	1	0	0	0	0	0
838	putative oxidoreductase subunit	630_gi|115251480	26	1	0	0	0	0	0
839	conserved hypothetical protein	630_gi|115251531	36	1	0	0	0	0	0
840	uracil-DNA glycosylase	630_gi|115251535	26	1	0	0	0	0	0
841	PyrR bifunctional protein includes: pyrimidine operon regulatory protein; uracil phosphoribosyltransferase	630_gi|115251652	20	1	0	0	0	0	0
842	tryptophanyl-tRNA synthetase	630_gi|115251661	38	1	0	0	0	0	0
843	putative ATP-binding protein	630_gi|115251691	18	1	0	0	0	0	0
844	oligopeptide ABC transporter, permease	630_gi|115251724	35	1	0	0	0	0	0

	protein								
845	probable polysaccharide deacetylase	630_gi|115251771	36	1	0	0	0	0	0
846	conserved hypothetical protein	630_gi|115251862	12	1	0	0	0	0	0
847	conserved hypothetical protein	630_gi|115252126	20	1	0	0	0	0	0
848	putative flavodoxin	630_gi|115252176	29	1	0	0	0	0	0
849	probable sigma54-dependent transcriptional regulator	630_gi|115252242	69	1	0	0	0	0	0
850	putative membrane protein	630_gi|115252285	36	1	0	0	0	0	0
851	hypoxanthine phosphoribosyltransferase	630_gi|115252288	20	1	0	0	0	0	0
852	probable GTP-binding protein	630_gi|115252356	22	1	0	0	0	0	0
853	GntR-family transcriptional regulator	630_gi|115252628	25	1	0	0	0	0	0
854	conserved hypothetical protein	630_gi|115252676	17	1	0	0	0	0	0
855	BASYS00041, 40158-39427 (CounterClockwise) Hypothetical Protein BASYS00041 Hypothetical protein	sta_BASYS00041 (115249290)	28	1	0	0	0	0	0
856	BASYS00235, 229993-230559 (Clockwise) Hypothetical Protein BASYS00235 signal peptide [*Clostridium difficile* QCD-66c26]	sta_BASYS00235	22	1	0	0	0	0	0
860	BASYS03583, 3840546-3841103 (Clockwise) Transcriptional Regulator PadR-Like Family	sta_BASYS03583	22	1	0	0	0	0	0
861	msbA, 3084615-3082825 (CounterClockwise) Hypothetical Protein msbA ABC transporter, ATP-binding/permease protein	stt_BASYS02922 (115251870)	67	1	0	0	0	0	0
869	ydgP, 1509321-1509890 (Clockwise) Electron Transport Complex Protein RnfG	sta_BASYS01412 115250170	20	0	1	0	0	0	0
870	conserved hypothetical protein	630_gi|115249704	20	1	0	0	0	0	0
872	putative transcriptional regulator	630_gi|115250151	21	0	1	0	0	0	0
875	putative membrane-associated neutral zinc metallopeptidase	630_gi|115251638	26	0	1	0	0	0	0
879	arginyl-tRNA synthetase	630_gi|115249727	65	1	0	0	0	0	0
881	putative multidrug efflux pump, membrane protein	630_gi|115251459	113	1	0	0	0	0	0
884	putative arsenate reductase	630_gi|115250823	14	1	0	0	0	0	0
885	conserved hypothetical protein	630_gi|115249717	22	1	0	0	0	0	0
894	aspartokinase	630_gi|115250358	45	1	0	0	0	0	0
897	ABC transporter, ATP-binding/permease protein	630_gi|115250108	65	1	0	0	0	0	0
899	putative sensor histidine kinase	630_gi|115251546	105	1	0	0	0	0	0
905	penicillin-binding protein	630_gi|115249798	97	0	1	0	0	0	0
907	putative gluconate permease	630_gi|115250781	43	1	0	0	0	0	0
916	DeoR-family transcriptional regulator (fatty acid and phospholipid biosynthesis regulator)	630_gi|115250210	21	1	0	0	0	0	0
917	PTS system, mannitol-specific IIa component	630_gi|115251387	15	1	0	0	0	0	0
919	BASYS03298, 3515475-3515182 (CounterClockwise) Hypothetical Protein BASYS03298 Hypothetical protein	sta_BASYS03298	11	1	0	0	0	0	1
927	putative AMP-binding protein	630_gi|115252027	53	1	0	0	0	0	0

928	putative histidyl-tRNA synthetase	630_gi|115251792	48	1	0	0	0	0	0
929	sta_BASYS02525-R	sta_BASYS02525	?	1	0	0	0	0	0

Table A3 – *Proteins identified by 2D reference mapping but not 1D GE LC-MS/MS.*

	Identification	Number	Mass	Mascot score			1D detected
				A	B-1	Tra5/5	
1	indolepyruvate oxidoreductase subunit	630_gi|115251433|emb|CAJ69266.1|	68929	195	307	102	- ABT
2	formate--tetrahydrofolate ligase	630_gi|115249735|emb|CAJ67552.1|	60347	162	206	132	- ABT
3	frc, 759341-760540 (Clockwise) Hypothetical Protein frc (isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase)	sta_BASYS00741	44491	87	184		- AB
4	thiH, 2730080-2731456 (Clockwise) Hypothetical Protein thiH (radical SAM-superfamily protein)	sta_BASYS02617	52653	94	216		- AB
5	gcvP, 1787909-1789366 (Clockwise) Hypothetical Protein gcvP (glycine cleavage system P protein)	stt_BASYS01732	53820	96		92	- AT
6	norV, 1741635-1744166 (Clockwise) Hypothetical Protein norV (putative nitric oxide reductase flavoprotein)	stt_BASYS01695	94959		178	122	- BT
7	glyS, 3040737-3038671 (CounterClockwise) Hypothetical Protein glyS (glycyl-tRNA synthetase beta chain)	stb_BASYS02649	78491		173	95	- BT
8	acpS, 64221-63841 (CounterClockwise) Hypothetical Protein acpS (holo-acyl-carrier protein synthase)	sta_BASYS00071	14473	56			- A
9	putative transcriptional regulator	630_gi|115252733|emb|CAJ70577.1|	50203	75			- A
10	putative flagellar hook-associated protein	630_gi|115249239|emb|CAJ67052.1|	48018	68			- A
11	ypdF, 2942563-2941484 (CounterClockwise) Hypothetical Protein ypdF (putative Xaa-Pro dipeptidase)	sta_BASYS02798	40623	82			- A
12	ptsI, 3427518-3425806 (CounterClockwise) Hypothetical Protein ptsI (phosphoenolpyruvate-protein phosphotransferase)	sta_BASYS03230	63296	182			- A
13	alanine racemase	630_gi|115252523|emb|CAJ70366.1|	43568		191		- B
14	fliY, 2449549-2448746 (CounterClockwise) Hypothetical Protein fliY (probable amino-acid ABC transporter,substrate-binding protein)	stb_BASYS02377	30120		147		- B
15	glycine/sarcosine/betaine reductase complex component C beta subunit	630_gi|115251404|emb|CAJ69236.1|	55464		130		- B
16	atpD, 3370184-3368406 (CounterClockwise) Hypothetical Protein atpD (V-type sodium ATP synthase subunit A)	stb_BASYS03178	65694		266		- B
17	BASYS00871, 876652-878778 (Clockwise) Hypothetical Protein BASYS00871	stt_BASYS00871	77451			172	- T
18	atpA, 3180789-3179011 (CounterClockwise) V-type ATP synthase alpha chain	stt_BASYS03006	65712			155	- T
19	rpsA, 1074685-1075959 (Clockwise)	stt_BASYS01053	48455			168	-

	Hypothetical Protein rpsA (putative 30S ribosomal protein S1)						T
20	prs, 3916457-3915507 (CounterClockwise) Hypothetical Protein prs (ribose-phosphate pyrophosphokinase)	stt_BASYS03658	34548			82	- T
21	ycgM, 996287-997189 (Clockwise) Hypothetical Protein ycgM (putative hydrolase)	stt_BASYS00983	33967			116	- T
22	efp, 1322094-1322666 (Clockwise) Hypothetical Protein efp (elongation factor P)	stt_BASYS01289	21572			66	- T
23	pgi, 3620557-3619208 (CounterClockwise) Hypothetical Protein pgi	stt_BASYS03385	50663			83	- T
24	aroA, 1980984-1981985 (Clockwise) Hypothetical Protein aroA (3--phosphoshikimate 1----carboxyvinyltransferase)	stt_BASYS01918	38093			99	- T
25	yjeK, 2408425-2407157 (CounterClockwise) Hypothetical Protein yjeK (L-lysine 2,3-aminomutase)	stt_BASYS02321	49419			92	- T
26	ATIC, 997288-998463 (Clockwise) Bifunctional purine biosynthesis protein PURH (putative formyltransferase)	stt_BASYS00984	43870			149	- T

Table A4 – *Proteins with significantly differing expression between the three strains as identified by correlation analysis using the SameSpots software with anova probability scores, Fold differences and the normalised volumes in each strain.*

Rank	Anova (p)	Fold	Notes	Average Normalised Volumes		
				B-1	A	Tra5/5
1	1.185e-006	43.7	ID: 105	0.124	5.414	0.488
2	0.004	33.4	ID: 002	0.076	0.800	2.539
3	1.506e-008	27.0	ID: 041	7.465	0.820	0.276
4	1.493e-008	22.6	ID: 270	2.268	0.128	0.100
5	3.865e-005	22.4	ID: 345	3.817	0.171	0.316
6	8.095e-004	21.1	ID: 351	3.918	0.186	0.577
7	7.514e-007	19.1	ID: 328	0.205	3.450	0.181
8	2.581e-004	17.8	ID: 360	0.522	0.111	1.970
9	1.271e-005	16.7	ID: 283	0.426	7.130	1.307
10	0.002	14.3	ID: 354	4.714	0.514	0.330
11	0.002	12.2	ID: 043	3.757	0.308	0.850
12	0.002	12.2	ID: 043	3.748	0.307	0.850
13	3.359e-004	11.4	ID: 042	1.804	0.158	0.731
14	0.026	11.1	ID: 257	0.374	2.843	4.149
15	2.994e-004	10.4	ID: 325	0.101	0.300	1.045
16	2.291e-004	9.7	ID: 044	1.110	0.114	0.505
17	3.979e-006	9.7	ID: 231	4.021	1.321	0.416
18	2.321e-004	9.6	ID: 359	0.383	3.672	0.707
19	0.010	9.6	ID: 323	0.132	0.342	1.271
20	0.001	9.3	ID: 005	0.843	5.510	0.590
21	0.009	8.8	ID: 336	0.189	0.594	1.652
22	4.595e-004	8.7	ID: 419	3.076	0.354	0.921
23	8.066e-004	8.2	ID: 024	0.507	4.142	0.663
24	0.042	7.9	ID: 006	0.229	0.373	1.810
25	0.002	7.9	ID: 282	5.609	0.711	1.223
26	2.892e-005	7.7	ID: 046	2.890	4.112	0.536
27	0.025	7.6	ID: 220	0.485	0.466	3.551
28	0.014	7.5	ID: 423	0.258	0.389	1.936
29	2.424e-004	7.4	ID: 070	0.640	4.725	1.434
30	0.004	7.4	ID: 132	1.870	0.253	1.413
31	0.003	7.1	ID: 122	0.543	3.835	1.396
32	3.935e-004	6.9	ID: 097	3.766	0.627	0.545
33	5.689e-004	6.8	ID: 020	4.634	0.686	0.742
34	2.526e-004	6.7	ID: 184	3.448	0.803	0.511

Rank	Anova (p)	Fold	Notes	Average Normalised Volumes		
				B-1	A	Tra5/5
35	0.005	6.7	ID: 394	0.389	1.054	2.623
36	3.765e-004	6.7	ID: 288	2.411	0.442	0.362
37	1.768e-005	6.6	ID: 198	4.716	0.769	0.717
38	0.015	6.5	ID: 066	4.719	0.724	1.015
39	0.015	6.4	ID: 348	0.755	2.135	0.334
40	0.035	6.3	ID: 098	0.346	0.646	2.168
41	0.005	6.0	ID: 340	2.923	0.484	0.592
42	0.002	6.0	ID: 392	0.417	2.506	0.507
43	0.002	5.8	ID: 067	0.475	2.771	1.632
44	0.041	5.8	ID: 156	2.988	0.513	2.423
45	0.016	5.6	ID: 188	1.022	4.016	0.711
46	5.077e-005	5.6	ID: 403	3.427	1.038	0.611
47	0.020	5.6	ID: 093	1.454	5.309	0.954
48	0.033	5.5	ID: 034	1.364	0.569	3.148
49	0.003	5.3	ID: 338	0.754	4.022	1.536
50	1.378e-004	5.3	ID: 054	0.794	4.218	0.867
51	2.726e-004	5.3	ID: 061	0.613	3.218	1.250
52	0.003	5.0	ID: 237	0.320	1.612	0.750
53	0.026	5.0	ID: 168	0.802	0.653	3.240
54	0.035	4.9	ID: 026	1.187	0.240	0.973
55	0.002	4.7	ID: 253	2.885	0.614	0.945
56	6.130e-005	4.7	ID: 135	0.706	3.313	0.994
57	0.002	4.6	ID: 191	2.171	0.512	0.472
58	0.003	4.6	ID: 402	0.788	3.589	2.287
59	0.001	4.5	ID: 395	1.350	0.348	1.583
60	0.007	4.5	ID: 344	0.587	2.068	0.464
61	0.005	4.4	ID: 350	1.265	3.161	0.719
62	0.013	4.4	ID: 365	0.447	0.656	1.947
63	0.005	4.3	ID: 173	1.985	0.636	2.740
64	0.002	4.3	ID: 341	0.584	2.230	0.523
65	0.009	4.2	ID: 021	4.142	0.992	1.086
66	0.002	4.1	ID: 045	0.889	0.214	0.782
67	0.005	4.1	ID: 440	3.453	0.846	1.410
68	6.251e-004	3.8	ID: 421	2.836	2.635	0.752
69	0.019	3.7	ID: 126	2.155	0.576	0.824
70	0.017	3.7	ID: 387	0.504	0.943	1.880
71	0.016	3.7	ID: 106	1.126	2.165	0.581
72	0.003	3.7	ID: 193	1.432	1.814	0.488
73	0.049	3.7	ID: 072	1.315	2.001	0.543

Rank	Anova (p)	Fold	Notes	Average Normalised Volumes		
				B-1	A	Tra5/5
74	0.037	3.5	ID: 311	0.338	0.411	1.175
75	0.020	3.4	ID: 146	2.963	0.864	1.569
76	0.029	3.4	ID: 037	1.573	0.979	0.461
77	0.021	3.4	ID: 092	0.190	0.274	0.642
78	0.012	3.3	ID: 233	2.301	1.559	0.707
79	0.006	3.1	ID: 211	1.982	1.539	0.643
80	0.002	3.1	ID: 117	1.254	3.344	1.088
81	0.009	3.1	ID: 049	1.125	0.600	1.843
82	0.048	3.1	ID: 130	2.944	0.981	0.961
83	0.014	3.0	ID: 408	2.706	2.112	0.901
84	0.026	2.9	ID: 162	0.378	0.404	1.095
85	0.018	2.9	ID: 051	1.970	0.685	0.817
86	0.027	2.8	ID: 166	2.000	0.702	1.493
87	0.009	2.8	ID: 127	1.603	0.575	1.102
88	0.006	2.8	ID: 194	2.080	0.776	0.752
89	0.002	2.7	ID: 241	2.926	2.654	1.071
90	0.003	2.7	ID: 239	1.851	0.851	0.682
91	0.005	2.7	ID: 225	1.045	1.801	0.670
92	0.015	2.7	ID: 136	0.462	1.232	0.848
93	0.007	2.6	ID: 453	1.775	0.683	0.909
94	1.744e-007	2.5	ID: 205	1.613	1.560	0.653
95	0.030	2.5	ID: 003	0.970	1.715	0.697
96	0.037	2.4	ID: 399	0.737	1.481	0.613
97	0.030	2.4	ID: 017	0.313	0.361	0.752
98	0.015	2.3	ID: 251	0.886	2.081	1.004
100	0.010	2.3	ID: 235	1.497	0.638	0.717
101	0.012	2.3	ID: 295	0.738	0.433	1.014
102	0.035	2.3	ID: 389	1.780	0.765	1.105
103	0.040	2.3	ID: 058	0.724	1.675	1.203
104	0.008	2.3	ID: 385	1.891	0.823	1.229
105	0.013	2.3	ID: 062	0.466	0.774	1.066
106	0.012	2.3	ID: 100	0.658	1.484	1.007
107	0.005	2.2	ID: 401	2.298	1.310	1.058
108	0.030	2.1	ID: 103	1.200	1.351	0.631
109	0.041	2.0	ID: 086	1.756	1.782	0.901
110	0.013	2.0	ID: 185	0.933	0.945	1.827
190	0.050	3.8	ID: 302	0.305	0.327	1.144

Table A5 – *Proteins identified in strain A ProteoMiner treated broth extracts by 1D GE followed by LC-MS/MS analysis with identification number, predicted mass and number of unique peptides detected in each sample. Theoretical pI values are also included for proteins identified only in treated fractions*

The number of proteins of different types are show as follows; metabolic (m), surface and virulence (s/v), hypothetical (h), regulatory (R), transporter (t) and ribosomal (r).

	Identification	Number	Mass	No. unique peptides				
				E-beads		SE-beads		Crude
	224 proteins - m=145, s/v=5, r=43, h=11, t=3, R=17			1	2	1	2	1
	All samples							
1	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405	41 kDa	28	25	23	25	21
2	pyruvate-flavodoxin oxidoreductase	630_gi|115251733	129 kDa	21	19	35	27	20
3	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	630_gi|115249401	44 kDa	22	19	18	18	14
4	electron transfer flavoprotein alpha-subunit	630_gi|115249407	37 kDa	18	14	12	16	16
5	30S ribosomal protein S3	630_gi|115249083	30 kDa	6	5	15	15	2
6	elongation factor TU	630_gi|115249061	44 kDa	17	15	16	10	13
7	electron transfer flavoprotein beta-subunit	630_gi|115249406	29 kDa	9	10	10	6	9
8	30S ribosomal protein S7	630_gi|115249073	18 kDa	14	17	12	10	9
9	putative amino acid aminotransferase	630_gi|115252729	45 kDa	24	19	17	17	14
10	putative rubrerythrin	630_gi|115250565	20 kDa	8	6	10	7	7
11	50S ribosomal protein L18	630_gi|115249093	14 kDa	5	3	8	6	1
12	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249403	47 kDa	20	19	13	10	18
13	acetyl-CoA acetyltransferase	630_gi|115250080	41 kDa	11	12	16	15	18
14	50S ribosomal protein L4	630_gi|115249078	24 kDa	6	7	12	10	4
15	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249404	42 kDa	20	14	12	8	16
16	30S ribosomal protein S5	630_gi|115249094	18 kDa	7	8	12	10	6
17	30S ribosomal protein S4	630_gi|115249105	24 kDa	12	10	8	7	7
18	putative oxidoreductase, thiamine diP-binding subunit	630_gi|115249125	39 kDa	12	9	7	10	5
19	adhE, 3632084-3629430 (CounterClockwise) Hypothetical Protein adhE aldehyde-alcohol dehydrogenase	sta_BASYS03400 115252023	97 kDa	17	14	20	12	21
20	succinate-semialdehyde dehydrogenase NAD(P)+	630_gi|115251397	51 kDa	11	10	16	13	20
21	50S ribosomal protein L19	630_gi|115250291	13 kDa	6	5	9	7	4
22	50S ribosomal protein L24	630_gi|115249088	11 kDa	3	2	4	5	2
23	50S ribosomal protein L13	630_gi|115249112	16 kDa	2	2	8	6	1
24	butyryl-CoA dehydrogenase	630_gi|115250075	41 kDa	11	10	10	9	12
25	50S ribosomal protein L1	630_gi|115249066	25 kDa	12	10	9	9	4
26	eutG, 2931498-2930332 (CounterClockwise) Hypothetical Protein eutG NAD-dependent 4-hydroxybutyrate dehydrogenase	sta_BASYS02789 115251393	44 kDa	11	12	8	9	8
27	DNA-directed RNA polymerase beta' chain	630_gi|115249071	130 kDa	8	6	28	17	13
28	30S ribosomal protein S9	630_gi|115249113	15 kDa	8	8	9	9	6
29	50S ribosomal protein L17	630_gi|115249107	13 kDa	2	2	5	4	1
30	NAD-specific glutamate dehydrogenase	630_gi|115249189	46 kDa	9	7	11	11	18

31	50S ribosomal protein L5	630_gi|115249089	20 kDa	13	12	7	8	9
32	BASYS02800, 2945591-2944029 (CounterClockwise) Hypothetical Protein BASYS02800 glycine/sarcosine/betaine reductase complex component C beta subunit [*Clostridium difficile* 630].	sta_BASYS02800 115251404	56 kDa	6	5	14	11	7
33	rpsJ, 420934-421245 (Clockwise) Hypothetical Protein rpsJ 30S ribosomal protein S10	sta_BASYS00411 115249076	12 kDa	6	4	5	5	3
34	30S ribosomal protein S20	630_gi|115251527	10 kDa	5	3	2	2	2
35	DNA-directed RNA polymerase beta chain	630_gi|115249070	139 kDa	18	13	17	12	19
36366	cell surface protein	630_gi|115251764	72 kDa	6	4	19	13	3
	yqjB, 3369095-3367134 (CounterClockwise) Uncharacterized protein yqjB cell surface protein	sta_BASYS03185 115251764	72 kDa	0	0	1	1	0
37	30S ribosomal protein S17	630_gi|115249086	10 kDa	4	3	4	2	2
38	aspartate aminotransferase	630_gi|115250375	45 kDa	17	12	9	8	11
39	30S ribosomal protein S13	630_gi|115249103	14 kDa	12	9	7	5	4
40	50S ribosomal protein L14	630_gi|115249087	13 kDa	3	4	7	4	1
41	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184	68 kDa	12	8	12	11	13
42	50S ribosomal protein L2	630_gi|115249080	30 kDa	7	7	7	7	2
43	50S ribosomal protein L6	630_gi|115249092	20 kDa	4	4	6	5	5
44	50S ribosomal protein L27	630_gi|115250195	10 kDa	4	3	7	5	2
45	stage 0 sporulation protein A	630_gi|115250247	31 kDa	5	4	12	13	6
46	norV, 2106146-2108677 (Clockwise) Hypothetical Protein norV putative nitric oxide reductase flavoprotein	sta_BASYS02024 115250664	94 kDa	12	12	7	4	16
47	50S ribosomal protein L15	630_gi|115249096	16 kDa	8	4	7	6	4
48	glyS, 3040737-3038671 (CounterClockwise) Hypothetical Protein glyS glycyl-tRNA synthetase beta chain	sta_BASYS02895 115251485	78 kDa	8	4	20	13	8
49	30S ribosomal protein S8	630_gi|115249091	15 kDa	4	5	6	5	6
50	50S ribosomal protein L22	630_gi|115249082	12 kDa	7	7	6	6	4
51	50S ribosomal protein L3	630_gi|115249077	22 kDa	8	7	6	3	5
52	putative ruberythrin	630_gi|115250515	20 kDa	4	4	4	2	3
53	glyceraldehyde-3-phosphate dehydrogenase 2	630_gi|115252231	36 kDa	7	8	9	5	7
54	50S ribosomal protein L21	630_gi|115250193	11 kDa	5	4	5	4	4
55	30S ribosomal protein S11	630_gi|115249104	14 kDa	4	5	4	2	4
56	50S ribosomal protein L16	630_gi|115249084	16 kDa	2	2	4	2	1
57	(R)-2-hydroxyisocaproate dehydrogenase	630_gi|115249400	37 kDa	9	7	6	6	10
58	pflD, 4068843-4066432 (CounterClockwise) Hypothetical Protein pflD putative formate acetyltransferase	sta_BASYS03788 115252338	91 kDa	12	9	10	5	12
59	proS2, 389555-391087 (Clockwise) Prolyl-tRNA synthetase 2 putative dual-specificity prolyl/cysteinyl-tRNA synthetase	sta_BASYS00379 115249053	59 kDa	9	7	8	7	5
60	30S ribosomal protein S19	630_gi|115249081	11 kDa	1	1	4	3	1
61	glycine reductase complex component B alpha and beta subunits	630_gi|115251407	46 kDa	8	9	6	6	12
62	30S ribosomal protein S2	630_gi|115251194	27 kDa	5	4	9	6	4

63	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 cell surface protein (S-layer precursor protein)	sta_BASYS03269 115251846	80 kDa	1	1	2	0	32
64	groL, 567992-569620 (Clockwise) Hypothetical Protein groL 60 kDa chaperonin	sta_BASYS00550 115249204	58 kDa	16	15	2	2	18
65	indolepyruvate oxidoreductase subunit	630_gi|115251433	66 kDa	9	7	6	7	7
66	electron transfer flavoprotein beta-subunit	630_gi|115250076	29 kDa	8	7	5	3	7
67	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,methyltransferase subunit	630_gi|115249744	29 kDa	9	9	6	4	6
68	putative aminoacyl-histidine dipeptidase	630_gi|115249724	53 kDa	9	5	7	4	7
69	putative anaerobic nitric oxide reductase flavorubredoxin	630_gi|115250189	45 kDa	6	5	3	1	5
70	metallo beta-lactamase superfamily protein	630_gi|115250323	62 kDa	10	7	8	8	3
71	putative subunit of oxidoreductase	630_gi|115249127	20 kDa	4	4	5	1	4
72	activator of 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249402	29 kDa	7	5	3	2	6
73	electron transfer flavoprotein alpha-subunit	630_gi|115250077	36 kDa	12	6	8	7	6
74	phosphate butyryltransferase	630_gi|115249121	32 kDa	2	7	5	4	5
75	lpdA, 1127096-1128496 (Clockwise) Hypothetical Protein lpdA putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	sta_BASYS01068 115249740	50 kDa	4	4	12	13	1
76	conserved hypothetical protein	630_gi|115252289	45 kDa	6	6	5	3	5
77	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079	31 kDa	6	6	5	5	7
78	gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase	630_gi|115251396	55 kDa	3	2	5	4	13
79	putative translation elongation factor	630_gi|115249025	72 kDa	9	7	10	3	12
80	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC flagellin subunit	sta_BASYS00598 115249247	34 kDa	6	6	9	6	5
81	translation initiation factor IF-3	630_gi|115249701	19 kDa	3	3	9	8	3
82	putative glutamate synthase NADPH small chain	630_gi|115250578	50 kDa	9	6	6	3	7
83	pyruvate, phosphate dikinase	630_gi|115251462	97 kDa	10	6	10	5	8
84	PTS system, IIb component	630_gi|115252071	11 kDa	2	1	4	3	1
85	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, alpha subunit	630_gi|115249743	50 kDa	5	3	5	2	4
86	isoleucyl-tRNA synthetase	630_gi|115251669	120 kDa	6	5	12	9	6
87	glycine/sarcosine/betaine reductase complex component C alpha subunit	630_gi|115251403	41 kDa	7	4	2	2	13
88	threonine dehydratase catabolic	630_gi|115251567	43 kDa	4	4	5	4	10
89	translation elongation factor G	630_gi|115249074	76 kDa	6	7	2	3	13
90	ferredoxin	630_gi|115249124	8 kDa	4	3	5	5	3
91	50S ribosomal protein L10	630_gi|115249067	19 kDa	3	5	6	4	8
92	enolase	630_gi|115252227	46 kDa	9	7	3	3	5

93	tig, 3643244-3641901 (CounterClockwise) Hypothetical Protein tig trigger factor	stt_BASYS03407 115252362	51 kDa	2	2	8	5	5
94	putative electron transfer protein	630_gi|115252303	47 kDa	3	4	13	9	4
95	ribose-phosphate pyrophosphokinase	630_gi|115252575	34 kDa	5	4	4	3	4
96	pyruvate carboxylase	630_gi|115249024	129 kDa	3	3	12	7	4
97	ypdF, 2851729-2849936 (CounterClockwise) Hypothetical Protein ypdF peptidase	sta_BASYS02718 115251315	68 kDa	6	7	3	2	8
98	30S ribosomal protein S16	630_gi|115250287	10 kDa	3	2	4	3	2
100	formate--tetrahydrofolate ligase	630_gi|115249735	60 kDa	4	3	5	5	13
101	leucyl-tRNA synthetase	630_gi|115251575	92 kDa	10	8	6	4	9
102	putative GTP-binding protein	630_gi|115251683	50 kDa	8	8	4	4	1
103	ATP synthase alpha chain	630_gi|115252530	55 kDa	6	4	3	3	8
104	putative propanediol utilization protein	630_gi|115251734	20 kDa	4	3	4	3	5
105	putative 5-nitroimidazole reductase	630_gi|115250500	18 kDa	3	4	1	1	7
106	putative alanyl-tRNA synthetase	630_gi|115250316	98 kDa	7	7	6	5	6
108	putative RNA-binding protein	630_gi|115249151	21 kDa	3	3	6	5	2
109	50S ribosomal protein L20	630_gi|115249703	13 kDa	3	3	3	3	2
110	rubrerythrin	630_gi|115249842	21 kDa	4	4	2	1	4
112	succinyl-CoA:coenzyme A transferase	630_gi|115251398	56 kDa	5	3	6	4	11
113	probable peptidase	630_gi|115251664	40 kDa	6	4	2	2	4
114	DNA gyrase subunit A	630_gi|115249009	91 kDa	2	1	14	7	5
115	DNA-directed RNA polymerase alpha chain	630_gi|115249106	35 kDa	8	6	8	6	3
116	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872	59 kDa	5	3	6	5	7
117	threonyl-tRNA synthetase	630_gi|115249589	74 kDa	3	3	7	4	4
118	putative 2-hydroxyacyl-CoA dehydratase	630_gi|115250793	44 kDa	3	3	3	2	4
119	ATP synthase beta chain	630_gi|115252528	50 kDa	3	4	4	2	4
120	phosphoenolpyruvate-protein phosphotransferase	630_gi|115251808	63 kDa	5	5	3	2	8
121	50S ribosomal protein L29	630_gi|115249085	8 kDa	2	2	3	3	2
122	50S ribosomal protein L9	630_gi|115252722	17 kDa	1	2	7	7	2
123	proC, 4066107-4065304 (CounterClockwise) Hypothetical Protein proC pyrroline-5-carboxylate reductase	sta_BASYS03787 115252337	28 kDa	5	5	3	1	2
124	putative phosphate butyryltransferase	630_gi|115249732	33 kDa	3	4	4	4	6
126	valS, 4034578-4031912 (CounterClockwise) Valyl-tRNA synthetase valyl-tRNA synthetase	sta_BASYS03755 115252312	103 kDa	6	3	4	3	4
127	manganese-dependent inorganic pyrophosphatase	630_gi|115249342	59 kDa	7	7	1	1	7
128	4-hydroxybutyrate CoA transferase	630_gi|115251394	48 kDa	1	1	2	1	12
129	putative molybdenum-binding subunit of oxidoreductase	630_gi|115251153	82 kDa	0	1	9	8	3
131	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein	630_gi|115250698	92 kDa	4	2	10	6	3
132	glucosamine--fructose-6-phosphate aminotransferase isomerizing	630_gi|115249129	67 kDa	2	3	3	2	7
133	transcription termination factor Rho	630_gi|115252548	61 kDa	5	4	3	3	1
135	conserved hypothetical protein	630_gi|115252075	17 kDa	2	2	2	2	2
137	50S ribosomal protein L7/L12	630_gi|115249068	13 kDa	5	5	2	2	4
138	conserved hypothetical protein	630_gi|115249575	22 kDa	2	1	4	2	1

140	GTP-sensing transcriptional pleiotropic repressor	630_gi|115250309	29 kDa	3	3	4	4	5
141	30S ribosomal protein S6	630_gi|115252728	11 kDa	5	4	2	0	3
142	ABC transporter, substrate-binding lipoprotein	630_gi|115249887	36 kDa	4	3	4	5	6
143	ATP-dependent protease La	630_gi|115252357	90 kDa	2	2	4	4	3
144	putative dinitrogenase iron-molybdenum cofactor	630_gi|115250736	13 kDa	1	1	1	1	2
145	butyrate kinase	630_gi|115249122	39 kDa	2	2	1	0	11
148	putative ornithine cyclodeaminase	630_gi|115249560	36 kDa	4	3	3	2	2
149	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	630_gi|115249132	45 kDa	2	1	6	4	1
150	putative FAD-binding subunit of oxidoreductase	630_gi|115251155	30 kDa	1	1	7	5	1
151	proline reductase subunit proprotein	630_gi|115252300	68 kDa	2	2	0	1	9
152	BASYS00434, 433091-433369 (Clockwise) Hypothetical Protein BASYS00434 putative ribosomal protein	sta_BASYS00434 115249100	11 kDa	3	1	6	5	1
153	uracil phosphoribosyltransferase	630_gi|115252539	23 kDa	1	3	5	2	1
154	agaY, 769977-769051 (CounterClockwise) Hypothetical Protein agaY putative fructose-bisphosphate aldolase	sta_BASYS00749 115249409	33 kDa	1	3	4	2	6
155	50S ribosomal protein L11	630_gi|115249065	15 kDa	1	1	2	2	1
156	tellurium resistance protein	630_gi|115250846	21 kDa	1	0	7	7	1
157	anaerobic ribonucleoside-triphosphate reductase	630_gi|115249116	89 kDa	4	2	4	2	3
158	putative NUDIX-family hydrolase	630_gi|115249807	17 kDa	2	1	3	2	3
159	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	630_gi|115249733	69 kDa	2	1	2	2	3
161	glycogen synthase	630_gi|115249898	56 kDa	4	4	1	1	1
162	conserved hypothetical protein	630_gi|115249288	15 kDa	3	2	3	2	3
163	putative dehydrogenase, electron transfer subunit	630_gi|115250577	33 kDa	3	4	1	0	4
164	50S ribosomal protein L23	630_gi|115249079	11 kDa	1	0	3	2	1
167	30S ribosomal protein S18	630_gi|115252726	9 kDa	3	2	1	1	0
168	glycine reductase complex component B gamma subunit	630_gi|115251405	47 kDa	3	1	1	1	6
172	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, small subunit	630_gi|115249742	34 kDa	3	2	2	3	2
173	S-adenosylmethionine synthetase	630_gi|115249139	44 kDa	2	1	3	4	1
174	trehalose-6-phosphate hydrolase	630_gi|115252145	66 kDa	2	3	2	1	3
175	putative DNA-binding protein	630_gi|115249180	23 kDa	2	1	4	3	2
176	gatZ, 588972-590237 (Clockwise) Hypothetical putative sugar-phosphate kinase	sta_BASYS00566 115249217	48 kDa	1	1	2	2	1
177	lysyl-tRNA synthetase	630_gi|115252615	58 kDa	3	2	2	2	5
178	putative universal stress protein	630_gi|115249829	15 kDa	3	3	1	2	3
181	tellurium resistance protein	630_gi|115250676	22 kDa	1	2	3	3	3
182	pyruvate kinase	630_gi|115252454	63 kDa	2	1	2	2	6
183	putative indolepyruvate oxidoreductase subunit	630_gi|115251432	21 kDa	2	1	2	2	2
184	putative pyridine nucleotide-disulfide oxidoreductase	630_gi|115250843	59 kDa	4	3	2	1	3

185	paaF, 1443446-1444243 (Clockwise) Hypothetical Protein paaF 3-hydroxybutyryl-CoA dehydratase	sta_BASYS01345 115250078	28 kDa	4	2	2	1	4
187	SpoIIIJ-associated protein	630_gi|115252741	24 kDa	3	3	2	3	2
188	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252652	119 kDa	3	3	2	3	4
190	V-type sodium ATP synthase subunit A	630_gi|115252012	66 kDa	3	3	0	1	2
191	putative transcriptional regulator	630_gi|115250535)	23 kDa	2	2	2	2	2
192	ribosome recycling factor	630_gi|115251191	21 kDa	1	1	2	2	3
193	RecA protein (recombinase A)	630_gi|115250364	37 kDa	1	1	2	2	1
197	cysteinyl-tRNA synthetase	630_gi|115249055	54 kDa	1	1	5	3	2
198	putative iron-only hydrogenase, catalytic subunit	630_gi|115252467	65 kDa	2	2	2	1	2
199	putative ATP-binding protein	630_gi|115250774	29 kDa	2	3	1	1	2
200	oligopeptide ABC transporter, substrate-binding protein	630_gi|115251723	58 kDa	1	0	2	1	6
202	GntR-family transcriptional regulator	630_gi|115250423	24 kDa	1	0	3	3	1
203	GntR-family transcriptional regulator	630_gi|115252628	25 kDa	2	2	2	2	2
204	DNA-binding protein HU	630_gi|115252557	10 kDa	3	1	2	1	2
205	gapA, 2263948-2264964 (Clockwise) Hypothetical Protein gapA glyceraldehyde-3-phosphate dehydrogenase	sta_BASYS02177 115250813	37 kDa	2	1	5	1	1
206	methionyl-tRNA synthetase	630_gi|115252601	74 kDa	3	3	1	0	4
207	3-oxoacyl-acyl-carrier-protein synthase II	630_gi|115250217	44 kDa	2	4	0	1	1
208	transketolase	630_gi|115251376	32 kDa	0	3	1	1	4
209	BASYS01830, 1919083-1918436 (CounterClockwise) Hypothetical Protein CTC Conserved hypothetical protein	sta_BASYS01830 115250551	24 kDa	1	3	1	1	3
213	conserved hypothetical protein	630_gi|115251617	13 kDa	3	1	2	1	2
214	putative 30S ribosomal protein S1	630_gi|115250007	48 kDa	4	2	3	0	1
215	conserved hypothetical protein	630_gi|115251576	13 kDa	1	1	1	1	1
219	conserved hypothetical protein	630_gi|115251019	48 kDa	2	1	1	1	1
220	elongation factor Ts	630_gi|115251193	33 kDa	1	1	1	1	5
221	uridylate kinase	630_gi|115251192	26 kDa	1	1	1	1	2
222	putative peptidase	630_gi|115252582	40 kDa	2	1	1	0	5
223	conserved hypothetical protein	630_gi|115250343	11 kDa	1	0	1	1	1
226	sigma-54-dependent transcriptional activator	630_gi|115252301	67 kDa	1	1	4	3	1
227	BASYS00907, 938664-938939 (Clockwise) Hypothetical Protein BASYS00907 Conserved hypothetical protein	sta_BASYS00907 115249602	10 kDa	1	1	1	1	2
230	DNA mismatch repair protein	630_gi|115251023	109 kDa	2	1	1	1	2
231	L-aspartate-beta-decarboxylase	630_gi|115251568	62 kDa	2	3	1	0	2
232	putative ATP/GTP-binding protein	630_gi|115249123	25 kDa	1	1	1	2	1
233	nitrilase (carbon-nitrogen hydrolase)	630_gi|115251895	34 kDa	2	2	1	1	0
243	ATP-dependent nuclease subunit B	630_gi|115250061	134 kDa	1	1	2	1	2
244	seryl-tRNA synthetase	630_gi|115249017	49 kDa	1	1	1	1	1
245	excinuclease ABC subunit A	630_gi|115252471	105 kDa	2	0	2	1	2
252	putative phosphoglucomutase	630_gi|115250617	64 kDa	2	2	1	1	2
253	ATP-dependent Clp protease	630_gi|115249029	91 kDa	2	1	3	0	2
254	putative ATP-dependent RNA helicase	630_gi|115249778	61 kDa	1	2	2	2	0
255	UDP-glucose 4-epimerase	630_gi|115251765	37 kDa	1	3	1	0	2
258	formylglycinamidine ribonucleotide synthetase	630_gi|115249233	141 kDa	1	0	1	0	5
269	anti-sigma F factor antagonist	630_gi|115249787	13 kDa	1	1	1	0	1

278	amidophosphoribosyltransferase	630_gi|115249228	51 kDa	2	1	1	1	1
279	cell division protein	630_gi|115251697	41 kDa	1	1	0	1	2
280	transcription antitermination protein	630_gi|115249064	20 kDa	1	1	1	1	1
281	trans-2-enoyl-ACP reductase	630_gi|115250213	33 kDa	1	1	2	1	1
284	proline reductase	630_gi|115252298	26 kDa	1	0	1	1	1
287	putative translation inhibitor endoribonuclease	630_gi|115251566	14 kDa	1	1	1	0	1
296	aspS, 3407936-3406149 (CounterClockwise) Hypothetical Protein asps putative aspartyl-tRNA synthetase	sta_BASYS03216 115251791	67 kDa	0	1	2	0	3
314	aspartate-semialdehyde dehydrogenase	630_gi|115252281	37 kDa	3	1	1	0	1
316	putative tRNA modification GTPase	630_gi|115252740	51 kDa	3	1	1	1	0
319	putative D-tyrosyl-tRNA protein	630_gi|115251795	16 kDa	1	1	1	0	1
320	phosphoribosylaminoimidazole carboxylase catalytic subunit	630_gi|115249226	17 kDa	0	1	1	1	1
332	cyclopropane-fatty-acyl-phospholipid synthase	630_gi|115249187	46 kDa	1	0	1	0	1
342	putative oxidoreductase, NAD/FAD binding subunit	630_gi|115249186	46 kDa	1	0	0	1	2
371	3-oxoacyl-acyl-carrier protein reductase	630_gi|115250215	27 kDa	1	1	1	1	1
405	aspartate aminotransferase	630_gi|115249115	44 kDa	0	1	1	0	1
441	conserved hypothetical protein	630_gi|115249281	54 kDa	1	0	0	1	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	111 proteins m=77, s/v=1, r=0, h=16, t=5, R=12			1	2	1	2	1
	Ebeads + crude							
111	asparaginyl-tRNA synthetase	630_gi|115251299	54 kDa	5	5	0	0	5
125	putative proline racemase	630_gi|115252294	36 kDa	4	5	0	0	6
130	pheT, 1093175-1095568 (Clockwise) Phenylalanyl-tRNA synthetase beta chain phenylalanyl-tRNA synthetase beta chain	sta_BASYS01042 115249716	89 kDa	7	6	0	0	10
136	thioredoxin reductase	630_gi|115251409	34 kDa	2	2	0	0	6
165	GMP synthase glutamine-hydrolyzing	630_gi|115249206	57 kDa	6	2	0	0	6
169	formate acetyltransferase	630_gi|115249776	84 kDa	5	4	0	0	5
170	ybgG, 3687856-3685256 (CounterClockwise) Hypothetical Protein ybgG methionyl-tRNA formyltransferase	sta_BASYS03452 115251640	100 kDa	7	3	0	0	9
179	chaperone protein (heat shock protein)	630_gi|115249282	75 kDa	3	2	0	0	9
180	ribonuclease R	630_gi|115252220	82 kDa	3	5	0	0	1
186	toxin A	630_gi|115249677	308 kDa	3	2	0	0	10
189	putative aminopeptidase	630_gi|115250096	51 kDa	1	2	0	0	1
195	AsnC-family transcriptional regulator	630_gi|115252605	16 kDa	5	2	0	0	4
196	hypothetical protein	630_gi|115249824	14 kDa	1	1	0	0	1
201	single-strand binding protein	630_gi|115252727	16 kDa	3	4	0	0	4
210	amino acid ABC transporter, substrate-binding protein	630_gi|115250820	30 kDa	2	2	0	0	6
212	BASYS01390, 1486311-1487180 (Clockwise) Hypothetical Protein BASYS01390 nitroreductase-family protein	sta_BASYS01390 115250155	33 kDa	2	0	0	0	6
216	ATP-dependent Clp protease ATP-binding subunit	630_gi|115252360	46 kDa	3	3	0	0	3
224	ferritin	630_gi|115251248	20 kDa	1	1	0	0	1
228	ribulose-phosphate 3-epimerase	630_gi|115251631	24 kDa	2	2	0	0	2
234	tellurium resistance protein	630_gi|115250677	23 kDa	2	3	0	0	3

235	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit	630_gi|115249745	77 kDa	2	2	0	0	2
236	translation initiation factor IF-2	630_gi|115250345	70 kDa	1	1	0	0	3
238	LysR-family transcriptional regulator	630_gi|115250103	35 kDa	1	3	0	0	3
239	dihydrodipicolinate synthase	630_gi|115252280	32 kDa	0	2	0	0	4
246	putative phosphomannomutase/phosphoglycerate mutase	630_gi|115251833	64 kDa	2	3	0	0	2
247	putative amidohydrolase	630_gi|115252278	42 kDa	2	1	0	0	2
248	putative serine hydroxymethyltransferase	630_gi|115251778	46 kDa	1	1	0	0	1
250	ABC transporter, ATP-binding protein	630_gi|115251122	62 kDa	1	0	0	0	6
257	anti-sigma F factor	630_gi|115249788	16 kDa	2	1	0	0	2
259	putative nitroreductase	630_gi|115251628	25 kDa	1	1	0	0	5
260	putative transaldolase	630_gi|115251384	23 kDa	1	1	0	0	4
262	conserved hypothetical protein	630_gi|115250197	14 kDa	1	1	0	0	2
263	putative cold shock protein	630_gi|115250391	7 kDa	1	1	0	0	2
264	polyribonucleotide nucleotidyltransferase	630_gi|115250354	78 kDa	2	0	0	0	4
267	putative glycine cleavage system H protein	630_gi|115249746	14 kDa	2	2	0	0	2
268	single-stranded DNA binding protein	630_gi|115252292	25 kDa	1	3	0	0	3
271	putative cyclase	630_gi|115251016	24 kDa	1	1	0	0	4
273	dihydrodipicolinate synthase	630_gi|115252282	32 kDa	0	2	0	0	3
275	NifU-like protein	630_gi|115250314	16 kDa	1	0	0	0	3
283	putative L-asparaginase	630_gi|115251569	36 kDa	1	4	0	0	2
285	putative phosphatase	630_gi|115251616	60 kDa	1	1	0	0	1
286	iron-dependent hydrogenase	630_gi|115249907	56 kDa	1	1	0	0	2
288	4-aminobutyrate aminotransferase	630_gi|115251211	48 kDa	1	1	0	0	1
289	putative decarboxylase	630_gi|115251040	12 kDa	2	0	0	0	4
294	probable hydrolase	630_gi|115251629	24 kDa	2	2	0	0	1
295	conserved hypothetical protein	630_gi|115252608	28 kDa	2	2	0	0	1
297	septum site-determining protein (cell division inhibitor)	630_gi|115250182	29 kDa	2	1	0	0	2
298	dipicolinate synthase, B chain	630_gi|115252025	21 kDa	1	2	0	0	2
299	conserved hypothetical protein	630_gi|115252645	33 kDa	2	2	0	0	1
300	ATP-dependent Clp protease proteolytic subunit	630_gi|115252361	21 kDa	1	1	0	0	3
301	radical SAM-superfamily protein	630_gi|115250190	71 kDa	1	1	0	0	2
302	conserved hypothetical protein	630_gi|115250191	27 kDa	2	1	0	0	1
303	glycyl-tRNA synthetase alpha chain	630_gi|115251486	34 kDa	1	1	0	0	1
304	conserved hypothetical protein	630_gi|115252525	17 kDa	3	0	0	0	2
306	ypdF, 2942563-2941484 (CounterClockwise) Hypothetical Protein ypdF putative Xaa-Pro dipeptidase	sta_BASYS02798 115251402	40 kDa	5	1	0	0	0
307	putative acetyltransferase	630_gi|115249853	19 kDa	1	0	0	0	3
317	putative 1-(5-phosphoribosyl)-5-(5-phosphoribosylamino)methylidene amino imidazole-4-carboxamide isomerase	630_gi|115250593	27 kDa	2	1	0	0	2
315	hcr, 1135233-1137227 (Clockwise) Hypothetical Protein Mbar A putative iron-sulfur protein	sta_BASYS01075 115249747	74 kDa	1	1	0	0	2
322	putative transcriptional repressor	630_gi|115252578	32 kDa	1	3	0	0	2
323	conserved hypothetical protein	630_gi|115249812	26 kDa	1	1	0	0	2
324	ABC transporter, ATP-binding protein	630_gi|115252673	27 kDa	1	1	0	0	2
325	putative metal dependent phosphohydrolase	630_gi|115251577	22 kDa	2	1	0	0	1
326	electron transfer flavoprotein beta-subunit	630_gi|115249821	29 kDa	1	1	0	0	1

327	putative phosphoribosyltransferase	630_gi|115251742	20 kDa	1	1	0	0	1
328	L-serine dehydratase	630_gi|115252279	43 kDa	1	1	0	0	1
329	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252654	119 kDa	1	1	0	0	1
330	putative selenocysteine lyase	630_gi|115252735	42 kDa	1	1	0	0	1
333	putative aliphatic sulfonates ABC transporter,substrate-binding lipoprotein	630_gi|115250525	36 kDa	0	1	0	0	3
338	conserved hypothetical protein	630_gi|115250029	32 kDa	1	2	0	0	1
339	threonine synthase	630_gi|115251172	56 kDa	1	1	0	0	2
340	proline iminopeptidase	630_gi|115252096	34 kDa	1	1	0	0	2
344	orotate phosphoribosyltransferase	630_gi|115249197	21 kDa	1	1	0	0	1
345	putative signal recognition particle	630_gi|115250284	47 kDa	1	1	0	0	1
346	putative NADPH-dependent FMN reductase	630_gi|115250716	23 kDa	1	1	0	0	1
347	putative carbonic anhydrase	630_gi|115251267	24 kDa	1	1	0	0	1
348	cell-division initiation protein	630_gi|115251670	20 kDa	1	1	0	0	1
349	rubrerythrin	630_gi|115251898	22 kDa	1	1	0	0	1
350	BASYS00737, 756750-756154 (CounterClockwise) Deoxyribonuclease TatD Family Radical SAM-superfamily protein	sta_BASYS00737 115249398	23 kDa	1	1	0	0	1
362	putative tellurite resistance protein	630_gi|115250680	44 kDa	1	0	0	0	2
363	10 kDa chaperonin	630_gi|115249203	10 kDa	1	0	0	0	1
375	conserved hypothetical protein	630_gi|115250440	24 kDa	1	1	0	0	1
376	radical SAM-superfamily protein	630_gi|115251209	52 kDa	1	1	0	0	1
377	two-component response regulator	630_gi|115251774	27 kDa	1	1	0	0	1
378	conserved hypothetical protein	630_gi|115252731	10 kDa	2	0	0	0	2
380	glutamyl-tRNA synthetase	630_gi|115249054	57 kDa	1	0	0	0	2
381	hypothetical protein	630_gi|115249604	92 kDa	1	0	0	0	2
382	conserved hypothetical protein	630_gi|115252618	28 kDa	1	0	0	0	2
385	phenylalanyl-tRNA synthetase alpha chain	630_gi|115249715	38 kDa	1	0	0	0	1
386	transketolase	630_gi|115251377	29 kDa	1	0	0	0	1
403	UDP-N-acetylmuramoylalanine--D-glutamate ligase	630_gi|115251704	50 kDa	1	1	0	0	1
404	inosine-uridine preferring nucleoside hydrolase	630_gi|115250725	36 kDa	1	1	0	0	1
407	putative nitrite and sulfite reductase subunit	630_gi|115250842	34 kDa	0	1	0	0	2
408	conserved hypothetical protein	630_gi|115249350	16 kDa	1	0	0	0	2
411	hydroxylamine reductase	630_gi|115251221	59 kDa	1	0	0	0	1
412	guanylate kinase	630_gi|115251645	23 kDa	1	0	0	0	1
413	ferredoxin	630_gi|115252670	6 kDa	1	0	0	0	1
414	BASYS00311, 304950-305525 (Clockwise) Hypothetical Protein BASYS00311 putative exported protein	sta_BASYS00311 115252734	21 kDa	1	0	0	0	1
416	(3R)-hydroxymyristoyl-acyl carrier protein dehydratase	630_gi|115249137	16 kDa	1	1	0	0	0
417	delta-aminolevulinic acid dehydratase	630_gi|115252479	36 kDa	1	1	0	0	0
444	putative phosphoribosylaminoimidazole-succinocarboxamide synthase	630_gi|115249503	25 kDa	0	1	0	0	1
445	probable amino-acid ABC transporter,substrate-binding protein	630_gi|115251230	30 kDa	0	1	0	0	1
446	putative pyruvate formate-lyase 3 activating enzyme	630_gi|115252339	34 kDa	0	1	0	0	1
447	DNA polymerase III, beta chain	630_gi|115249005	46 kDa	1	0	0	0	1
448	xanthine phosphoribosyltransferase	630_gi|115251385	21 kDa	1	0	0	0	1
449	S-ribosylhomocysteinase	630_gi|115252662	17 kDa	1	0	0	0	1
462	putative ATP-binding protein	630_gi|115250457	25 kDa	1	0	0	0	1

486	conserved hypothetical protein	630_gi|115250280	17 kDa	1	0	0	0	1
498	phosphopantetheine adenylyltransferase	630_gi|115251613	18 kDa	0	1	0	0	1
511	putative cation transport protein	630_gi|115249713	24 kDa	1	0	0	0	1
512	hypothetical protein	630_gi|115252039	74 kDa	1	0	0	0	1
572	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	630_gi|115251182	38 kDa	1	0	0	0	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	46 proteins - m=27, s/v=0, r=0, h=13, t=2, R=4			1	2	1	2	1
	Ebeads only							
331	putative thiamine biosynthesis protein	630_gi|115249771	44 kDa 6.97	3	2	0	0	0
353	conserved hypothetical protein	630_gi|115252583	51 kDa 4.45	2	2	0	0	0
392	SsrA-binding protein	630_gi|115252213	19 kDa 9.64	1	1	0	0	0
437	conserved hypothetical protein	630_gi|115252126	20 kDa 5.97	0	2	0	0	0
442	conserved hypothetical protein	630_gi|115251612	20 kDa 4.75	1	2	0	0	0
443	putative phosphoserine phosphatase	630_gi|115249318	28 kDa 5.06	1	2	0	0	0
450	conserved hypothetical protein	630_gi|115251509	29 kDa 9.43	1	1	0	0	0
472	putative aromatic amino acid aminotransferase	630_gi|115251881	45 kDa 5.01	2	0	0	0	0
475	ferric uptake regulation protein	630_gi|115250321	18 kDa 5.22	2	0	0	0	0
480	putative transcriptional regulator	630_gi|115252733	51 kDa 5.48	0	1	0	0	0
482	glycerol dehydratase	630_gi|115250149	89 kDa 5.64	1	0	0	0	0
487	metallo-beta-lactamase superfamily protein	630_gi|115249353	35 kDa 5.38	1	1	0	0	0
488	conserved hypothetical protein	630_gi|115251511	30 kDa 5.20	1	1	0	0	0
504	aspartate--ammonia ligase	630_gi|115251746	39 kDa 4.86	0	2	0	0	0
540	conserved hypothetical protein	630_gi|115249592	22 kDa 5.11	0	1	0	0	0
541	putative radical SAM superfamily lipoprotein	630_gi|115251281	33 kDa 8.96	0	1	0	0	0
543	BASYS03624, 3892755-3891991 (CounterClockwise) Conserved Hypothetical Protein -	sta_BASYS03624	30 kDa 9.11	0	1	0	0	0
545	transcription elongation protein	630_gi|115250342	44 kDa 4.59	1	0	0	0	0
546	conserved hypothetical protein	630_gi|115250366	70 kDa 5.19	1	0	0	0	0
547	PTS system, lichenan-specific IIa component	630_gi|115251933	12 kDa 5.18	1	0	0	0	0

563	geranyltranstransferase	630_gi|115250238	33 kDa 5.06	0	1	0	0	0
565	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252465	18 kDa 6.29	1	0	0	0	0
614	aspartokinase	630_gi|115250358	45 kD 4.95	0	1	0	0	0
615	putative sigma 54 modulation protein	630_gi|115251497	22 kDa 4.99	0	1	0	0	0
616	glucokinase	630_gi|115251513	34 kDa 5.44	0	1	0	0	0
617	glycerol-3-phosphate dehydrogenase NAD(P)+	630_gi|115251681	37 kDa 6.91	0	1	0	0	0
618	conserved hypothetical protein	630_gi|115252038	35 kDa 6.79	0	1	0	0	0
619	prolyl-tRNA synthetase	630_gi|115249052	64 kDa 5.11	1	0	0	0	0
620	putative amidohydrolase	630_gi|115249796	41 kDa 5.16	1	0	0	0	0
621	exonuclease subunit C	630_gi|115250064	136 kDa 6.63	1	0	0	0	0
622	conserved hypothetical protein	630_gi|115250233	13 kDa 5.21	1	0	0	0	0
623	putative pyridoxine kinase	630_gi|115250389	31 kDa 4.53	1	0	0	0	0
624	tRNA pseudouridine synthase 2	630_gi|115250830	29 kDa 9.14	1	0	0	0	0
625	putative ethanolamine/propanediol utilization cobalamin adenosyltransferase	630_gi|115250966	30 kDa 8.99	1	0	0	0	0
626	1-deoxy-d-xylulose 5-phosphate reductoisomerase	630_gi|115251184	43 kDa 5.15	1	0	0	0	0
627	putative amidohydrolase	630_gi|115251748	41 kDa 5.53	1	0	0	0	0
628	putative esterase/halogenase	630_gi|115251917	30 kDa 5.11	1	0	0	0	0
629	ypdF, 3268469-3267330 (CounterClockwise) Hypothetical Protein ypdF -	sta_BASYS03098	43 kDa 5.36	1	0	0	0	0
635	BASYS02150, 2243043-2243318 (Clockwise) Hypothetical Protein BASYS02150	sta_BASYS02150	10 kDa 10.76	0	1	0	0	0
639	tRNA pseudouridine synthase 1	630_gi|115249111	27 kDa 8.99	0	1	0	0	0
654	conserved hypothetical protein	630_gi|115251206	9 kDa 7.70	1	0	0	0	0
679	probable esterase	630_gi|115251258	30 kDa 5.18	1	0	0	0	0
690	ABC transporter, ATP-binding protein	630_gi|115251001	25 kDa 7.78	1	0	0	0	0
700	putative sulfonate ABC transporter,solute-binding lipoprotein	630_gi|115251417	40 kDa 5.79	1	0	0	0	0
707	cobalt transport protein	630_gi|115249335	27 kDa 8.98	1	0	0	0	0

711	putative tRNA/rRNA methyltransferase	630_gi|115249058	27 kDa 6.60	1	0	0	0	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	19 proteins - m=15, s/v=1, r=1, h=3, t=0, R=0			1	2	1	2	1
	Ebeads and SE beads							
99	30S ribosomal protein S15	630_gi|115250352	10 kDa 10.19	7	7	5	2	0
107	adenylosuccinate lyase	630_gi|115250370	55 kDa 5.82	3	2	9	5	0
261	putative GTP pyrophosphokinase	630_gi|115251796	85 kDa 8.81	1	1	2	0	0
318	probable GTP-binding protein	630_gi|115252356	22 kDa 8.91	0	1	2	2	0
321	glycine cleavage system P protein	630_gi|115250699	54 kDa 5.55	1	1	1	1	0
372	glucose inhibited division protein A	630_gi|115252739	71 kDa 6.26	1	0	2	1	0
373	putative 16S rRNA processing protein	630_gi|115250289	20 kDa 5.07	1	1	0	2	0
387	signal recognition particle protein	630_gi|115250286	50 kDa 9.26	1	0	1	0	0
415	undecaprenyl pyrophosphate synthetase	630_gi|115251189	28 kDa 8.62	0	1	1	1	0
439	putative radical SAM family protein	630_gi|115251801	53 kDa 5.04	1	1	1	0	0
440	Spo0B-associated GTP-binding protein	630_gi|115250196	47 kDa 5.11	1	0	1	1	0
452	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	630_gi|115249595	53 kDa 6.44	1	0	0	1	0
455	hypothetical protein	630_gi|115252720	12 kDa 4.77	1	0	1	0	0
490	GTP-binding protein	630_gi|115251490	34 kDa 8.33	0	1	1	0	0
491	putative molybdopterin biosynthesis protein	630_gi|115251123	37 kDa 5.42	1	0	1	0	0
492	BASYS02918, 3060611-3059598 (CounterClockwise) Hypothetical Protein Aq conserved hypothetical protein	sta_BASYS02918 115251507	39 kDa 5.58	1	0	1	0	0
493	biotin carboxylase (acetyl-CoA carboxylase subunit A)	630_gi|115250984	51 kDa 5.37	0	1	1	0	0
510	peptidase T	630_gi|115250067	45 kDa 4.86	1	0	1	0	0
551	conserved hypothetical protein	630_gi|115250989	11 kDa 4.34	1	0	0	1	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	22 proteins m=14, s/v=2, r=0, h=4, t=0, R=2			1	2	1	2	1
	SE beads and crude							
134	BASYS03267, 3473846-3471975 (CounterClockwise) Hypothetical Protein BASYS03267 cell surface protein (putative S-layer protein precursor)	sta_BASYS03267 115251844	66 kDa	0	0	6	4	11

146	conserved hypothetical protein	630_gi|115249020	12 kDa	0	0	5	5	1
229	preprotein translocase SecA subunit	630_gi|115249152	102 kDa	0	0	4	1	3
256	BASYS00644, 660422-660628 (Clockwise) Hypothetical Protein LJ conserved hypothetical protein	sta_BASYS00644 115249289	8 kDa	0	0	2	2	1
270	putative thiol peroxidase (bacterioferritin comigratory protein)	630_gi|115250867	18 kDa	0	0	2	1	3
274	TetR-family transcriptional regulator	630_gi|115251558	35 kDa	0	0	1	1	1
282	porphyrin biosynthesis protein includes: uroporphyrinogen-III methyltransferase and uroporphyrinogen-III synthase	630_gi|115252480	56 kDa	0	0	2	2	1
290	glycyl-tRNA synthetase beta chain	630_gi|115251485	78 kDa	0	0	2	2	0
291	secA, 3476606-3474261 (CounterClockwise) Hypothetical Protein secA preprotein translocase SecA subunit	sta_BASYS03268 115251845	89 kDa	0	0	2	1	1
305	riboflavin biosynthesis protein	630_gi|115250350	35 kDa	0	0	3	2	0
309	thioredoxin	630_gi|115250733	12 kDa	0	0	1	0	3
341	ATP synthase subunit delta	630_gi|115252531	21 kDa	0	0	2	1	1
357	transketolase, thiamine disphosphate-binding subunit	630_gi|115252520	30 kDa	0	0	1	1	1
358	conserved hypothetical protein	630_gi|115251463	32 kDa	0	0	3	0	1
374	conserved hypothetical protein	630_gi|115251464	23 kDa	0	0	1	1	2
379	putative methylenetetrahydrofolate reductase	630_gi|115249739	32 kDa	0	0	1	1	1
406	two-component response regulator	630_gi|115250828	28 kDa	0	0	1	1	1
409	penicillin-binding protein	630_gi|115249798	97 kDa	0	0	1	0	2
410	tpiA, 3926737-3925994 (CounterClockwise) Hypothetical Protein tpiA triosephosphate isomerase	sta_BASYS03660 115252229	27 kDa	0	0	1	0	2
453	PyrR bifunctional protein includes: pyrimidine operon regulatory protein; uracil phosphoribosyltransferase	630_gi|115251652	20 kDa	0	0	1	0	1
454	cell surface protein (putative cell surface-associated cysteine protease)	630_gi|115251840	87 kDa	0	0	1	0	1
497	fibronectin-binding protein	630_gi|115251649	68 kDa	0	0	0	1	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	83 proteins m=34, s/v=9, r=7, h=15, t=2, R=16			1	2	1	2	1
	SE beads only							
	putative glycosyl transferase BASYS00343, 340979-339879	630_gi|115249022	45 kDa 7.66	0	0	6	5	0
139424	(CounterClockwise) Hypothetical Protein BASYS00343 putative glycosyl transferase	sta_BASYS00343 115249022	41 kDa	0	0	1	1	0
160	RNA polymerase sigma factor (sigma-43)	630_gi|115250496	44 kDa 5.12	0	0	7	6	0
166	putative subunit of oxidoreductase	630_gi|115249126	27 kDa 6.38	0	0	4	1	0
171	conserved hypothetical protein	630_gi|115250341	18 kDa 4.82	0	0	3	1	0
211	30S ribosomal protein S21	630_gi|115251500	7 kDa 10.78	0	0	3	3	0
217	ABC transporter, ATP-binding protein	630_gi|115251510	42 kDa 5.98	0	0	2	2	0
218	putative formate dehydrogenase	630_gi|115251232	83 kDa 8.71	0	0	3	2	0

225	rbfA, 1693204-1693584 (Clockwise) Hypothetical Protein rbfA ribosome-binding factor A	sta_BASYS01610 115250346	14 kDa 5.34	0	0	4	2	0
240	V-type sodium ATP synthase subunit D	630_gi|115252010	26 kDa 9.08	0	0	4	3	0
241	50S ribosomal protein L32	630_gi|115250209	6 kDa 9.84	0	0	1	1	0
242	putative RNA-binding protein	630_gi|115252556	10 kDa 9.44	0	0	1	1	0
249	DeoR-family transcriptional regulator (fatty acid and phospholipid biosynthesis regulator)	630_gi|115250210	21 kDa 9.11	0	0	3	3	0
265	BASYS02843, 2985540-2985163 (CounterClockwise) Hypothetical Protein BASYS02843 putative beta-lactamase repressor	sta_BASYS02843 115251442	15 kDa 6.91	0	0	3	2	0
266	putative membrane protein	630_gi|115250942	14 kDa 9.52	0	0	2	2	0
272	putative glycosyltransferase	630_gi|115250250	26 kDa 9.07	0	0	4	2	0
334	conserved hypothetical protein	630_gi|115252730	7 kDa 9.51	0	0	2	2	0
336	putative DNA topoisomerase	630_gi|115251327	81 kDa 9.43	0	0	1	1	0
351	putative capsular polysaccharide biosynthesis glycosyl transferase	630_gi|115251823	41 kDa 6.48	0	0	2	3	0
352	UDP-N-acetylglucosamine--N-acetylmuramyl-(penta peptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	630_gi|115251702	45 kDa 9.46	0	0	2	2	0
354	histidinol dehydrogenase	630_gi|115250622	47 kDa 5.46	0	0	2	2	0
359	putative competence protein	630_gi|115250449	46 kDa 6.35	0	0	2	1	0
360	putative histidinol-phosphate aminotransferase	630_gi|115250590	40 kDa 5.28	0	0	2	1	0
361	cell surface protein	630_gi|115251032	52 kDa 9.40	0	0	2	1	0
383	conserved hypothetical protein	630_gi|115249751	7 kDa 4.61	0	0	2	2	0
384	GntR-family transcriptional regulator	630_gi|115251077	14 kDa 5.11	0	0	2	1	0
388	50S ribosomal protein L30	630_gi|115249095	7 kDa 9.87	0	0	1	1	0
389	SOS regulatory protein	630_gi|115250977	24 kDa 5.38	0	0	1	1	0
390	conserved hypothetical protein	630_gi|115251876	53 kDa 6.19	0	0	1	1	0
418	translation initiation factor IF-1	630_gi|115249101	8 kDa 9.25	0	0	1	1	0
420	putative transcriptional regulator	630_gi|115250685	21 kDa 6.11	0	0	1	1	0
421	conserved hypothetical protein	630_gi|115251421	12 kDa 5.14	0	0	1	1	0

422	D-alanine--poly(phosphoribitol) ligase subunit 2 (D-alanyl carrier protein)	630_gi|115251904	9 kDa 4.05	0	0	1	1	0
423	conserved hypothetical protein	630_gi|115252658	36 kDa 6.21	0	0	1	1	0
425	BASYS03265, 3471128-3469293 (CounterClockwise) Hypothetical Protein BASYS03265 cell surface protein	sta_BASYS03265 115251842	67 kDa 5.30	0	0	1	1	0
456	NUDIX-family protein	630_gi|115251029	24 kDa 6.12	0	0	1	1	0
457	putative radical SAM superfamily protein	630_gi|115251502	50 kDa 7.01	0	0	1	1	0
458	conserved hypothetical protein	630_gi|115252542	39 kDa 6.88	0	0	1	1	0
463	electron transport complex protein	630_gi|115250168	48 kDa 6.30	0	0	3	0	0
479	putative DNA-binding protein	630_gi|115249855	12 kDa 9.05	0	0	1	1	0
494	RRF2-family transcriptional regulator	630_gi|115250312	16 kDa 7.60	0	0	1	1	0
495	stage IV sporulation protein A	630_gi|115251680	56 kDa 4.97	0	0	1	1	0
496	queuine tRNA-ribosyltransferase	630_gi|115251855	43 kDa 8.06	0	0	1	1	0
506	BirA bifunctional protein includes: biotin operon repressor and biotin--acetyl-CoA-carboxylase synthetase	630_gi|115252621	37 kDa 6.07	0	0	1	0	0
550	putative toxic anion resistance protein	630_gi|115251391	44 kDa 5.60	0	0	0	1	0
553	30S ribosomal protein S12	630_gi|115249072	16 kDa 11.09	0	0	1	0	0
554	putative flagellar protein	630_gi|115249264	8 kDa 9.57	0	0	1	0	0
555	putative DNA mismatch repair protein	630_gi|115249725	88 kDa 6.08	0	0	1	0	0
556	conserved hypothetical protein	630_gi|115250444	14 kDa 6.09	0	0	1	0	0
557	putative ferrous iron transport protein A	630_gi|115250789	8 kDa 9.82	0	0	1	0	0
558	stage V sporulation protein S	630_gi|115250981	9 kDa 7.89	0	0	1	0	0
559	conserved hypothetical protein	630_gi|115251395	10 kDa 6.07	0	0	1	0	0
560	ATP synthase epsilon chain (partial)	630_gi|115252527	9 kDa 6.57	0	0	1	0	0
561	csrA, 612768-612980 (Clockwise) Hypothetical Protein csrA carbon storage regulator	sta_BASYS00593 115249242	8 kDa 9.79	0	0	1	0	0
562	BASYS02498, 2595334-2595041 (CounterClockwise) Hypothetical Protein BASYS02498 hypothetical protein	sta_BASYS02498 115251100	12 kDa 5.36	0	0	1	0	0
573	putative tellurium resistance protein	630_gi|115250693	22 kDa 7.71	0	0	1	0	0

575	fabD, 1551856-1552806 (Clockwise) Hypothetical Protein fabD malonyl CoA-acyl carrier protein transacylase	sta_BASYS01457 115250214	34 kDa 5.25	0	0	1	0	0
576	L-lysine 2,3-aminomutase	630_gi|115251307	49 kDa 6.08	0	0	1	0	0
630	putative iron-sulfur protein	630_gi|115250509	48 kDa 8.27	0	0	0	1	0
631	conserved hypothetical protein	630_gi|115250815	159 kDa 7.63	0	0	0	1	0
632	putative ATP-dependent RNA helicase	630_gi|115251160	43 kDa 9.46	0	0	0	1	0
633	chaperone protein	630_gi|115251514	42 kDa 8.85	0	0	0	1	0
634	carbamoyl-phosphate synthase,pyrimidine-specific, small chain	630_gi|115252653	39 kDa 5.79	0	0	0	1	0
642	flagellar protein FliS	630_gi|115249243	16 kDa 9.07	0	0	1	0	0
643	penicillinase repressor	630_gi|115249484	15 kDa 6.62	0	0	1	0	0
644	fatty acid/phospholipid synthesis protein	630_gi|115250211	37 kDa 5.97	0	0	1	0	0
645	putative ribosomal protein	630_gi|115250344	11 kDa 9.62	0	0	1	0	0
646	cell surface protein (putative cell surface-associated cysteine protease)	630_gi|115250795	87 kDa 6.49	0	0	1	0	0
647	ribonuclease Ph	630_gi|115252364	50 kDa 6.99	0	0	1	0	0
648	dimethyladenosine transferase	630_gi|115252584	32 kDa 6.47	0	0	1	0	0
649	ydhO, 1504418-1505713 (Clockwise) Hypothetical Protein ydhO putative cell wall hydrolase	sta_BASYS01408 115250166	45 kDa 9.47	0	0	1	0	0
653	conserved hypothetical protein	630_gi|115252701	35 kDa 6.22	0	0	0	1	0
656	putative phosphatase	630_gi|115250093	31 kDa 5.11	0	0	0	1	0
660	BASYS02611, 2724090-2723929 (CounterClockwise) Hypothetical Protein BASYS02611	sta_BASYS02611 -	6 kDa 9.46	0	0	1	0	0
662	50S ribosomal protein L33	630_gi|115249062	6 kDa 9.91	0	0	0	1	0
667	putative oxidoreductase	630_gi|115250433	49 kDa 6.35	0	0	1	0	0
668	cell division protein	630_gi|115252622	74 kDa 5.61	0	0	1	0	0
671	anti-sigma-B factor antagonist	630_gi|115249012	12 kDa 6.08	0	0	1	0	0
683	putative helicase	630_gi|115250059	87 kDa 5.78	0	0	1	0	0
686	coenzyme A biosynthesis bifunctional protein	630_gi|115251643	44 kDa 7.59	0	0	1	0	0
703	ABC transporter, ATP-binding protein	630_gi|115249108	31 kDa 4.89	0	0	0	1	0

708	rod shape-determining protein	630_gi|115250177	38 kDa 4.99	0	0	1	0	0
709	hsdR, 3762409-3759074 (CounterClockwise) Hypothetical Protein hsdR	sta_BASYS03516 -	129 kDa 5.43	0	0	1	0	0
717	PadR-family transcriptional regulator	630_gi|115250186	13 kDa 9.08	0	0	0	1	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	138 proteins m=88, s/v=12, r=0, h=22, t=4, R=12			1	2	1	2	1
	Crude only							
147	chaperone protein	630_gi|115251515	66 kDa	0	0	0	0	10
237	60 kDa chaperonin	630_gi|115249204	58 kDa	1	1	0	0	2
251	acetate kinase	630_gi|115250207	43 kDa	0	0	0	0	5
276	cell surface protein (putative penicillin-binding protein)	630_gi|115250510	111 kDa	0	0	0	0	5
292	6-phosphofructokinase	630_gi|115252455	34 kDa	0	0	0	0	4
293	elongation factor P	630_gi|115250279	21 kDa	0	0	0	0	2
310	inosine-5'-monophosphate dehydrogenase	630_gi|115251390	55 kDa	0	0	0	0	5
311	conserved hypothetical protein	630_gi|115251601	68 kDa	0	0	0	0	5
312	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898 DNA repair protein	stt_BASYS02898 115251488	80 kDa	0	0	0	0	3
313	conserved hypothetical protein	630_gi|115252732	22 kDa	0	0	0	0	3
337	cell surface protein	630_gi|115249861	34 kDa	0	0	0	0	4
367	putative peptidyl-prolyl isomerase	630_gi|115250393	28 kDa	0	0	0	0	3
368	adenylate kinase	630_gi|115249098	24 kDa	0	0	0	0	4
369	transcription elongation factor	630_gi|115252616	18 kDa	0	0	0	0	4
370	butyrate kinase	630_gi|115251431	39 kDa	0	0	0	0	3
391	BASYS00898, 930141-930440 (Clockwise) Hypothetical Protein BASYS00898 -	sta_BASYS00898 -	11 kDa	0	0	0	0	1
393	hypothetical protein	630_gi|115251768	78 kDa	0	0	0	0	4
394	arginyl-tRNA synthetase	630_gi|115249727	65 kDa	0	0	0	0	3
395	putative histidinol-phosphate aminotransferase	630_gi|115251555	42 kDa	0	0	0	0	3
396	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	630_gi|115252228	56 kDa	0	0	0	0	3
397	conserved hypothetical protein	630_gi|115249879	33 kDa	0	0	0	0	2
398	uracil-DNA glycosylase	630_gi|115251535	26 kDa	0	0	0	0	2
399	putative alanine racemase	630_gi|115251674	27 kDa	0	0	0	0	2
400	glucose-6-phosphate isomerase	630_gi|115252341	51 kDa	0	0	0	0	2
401	putative tRNA binding protein	630_gi|115251499	17 kDa	0	0	0	0	2
427	dihydroorotate dehydrogenase, catalytic subunit	630_gi|115249196	32 kDa	0	0	0	0	4
428	clpB, 2575497-2572903 (CounterClockwise) Hypothetical Protein clpB chaperone	sta_BASYS02475 115251074	98 kDa	0	0	0	0	3
429	V-type sodium ATP synthase subunit B	630_gi|115252011	51 kDa	0	0	0	0	3
430	putative phosphoglucomutase	630_gi|115251373	56 kDa	0	0	0	0	3
431	ABC transporter, ATP-binding protein	630_gi|115249802	59 kDa	0	0	0	0	2
432	PTS system, IIb component	630_gi|115249296	18 kDa	0	0	0	0	2
433	peptidyl-prolyl cis-trans isomerase	630_gi|115249340	19 kDa	0	0	0	0	2

434	DNA polymerase I	630_gi|115250159	101 kDa	0	0	0	0	2
435	aspartokinase	630_gi|115251108	43 kDa	0	0	0	0	2
436	conserved hypothetical protein	630_gi|115251678	10 kDa	0	0	0	0	2
438	tellurium resistance protein	630_gi|115250675	21 kDa	0	0	0	0	1
464	dihydroorotate dehydrogenase electron transfer subunit	630_gi|115249195	26 kDa	0	0	0	0	2
465	conserved hypothetical protein	630_gi|115249410	13 kDa	0	0	0	0	2
466	putative carbon-nitrogen hydrolase	630_gi|115249501	31 kDa	0	0	0	0	2
467	RpiR-family transcriptional regulator	630_gi|115251102	33 kDa	0	0	0	0	2
468	putative glutamyl-aminopeptidase	630_gi|115251205	39 kDa	0	0	0	0	2
469	thioredoxin	630_gi|115251408	12 kDa	0	0	0	0	2
470	heat shock protein	630_gi|115251516	24 kDa	0	0	0	0	2
471	cell surface protein	630_gi|115251837	73 kDa	0	0	0	0	2
473	dihydrodipicolinate reductase	630_gi|115252283	27 kDa	0	0	0	0	2
474	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	630_gi|115252284	25 kDa	0	0	0	0	2
476	UDP-N-acetylenolpyruvoylglucosamine reductase	630_gi|115252462	33 kDa	0	0	0	0	2
478	HPr(Ser) kinase/phosphorylase	630_gi|115252469	36 kDa	0	0	0	0	1
483	conserved hypothetical protein	630_gi|115250701	12 kDa	0	0	0	0	1
499	conserved hypothetical protein	630_gi|115249349	195 kDa	0	0	0	0	2
500	cell surface protein (putative hemagglutinin/adhesin)	630_gi|115249532	167 kDa	0	0	0	0	2
501	conserved hypothetical protein	630_gi|115249797	34 kDa	0	0	0	0	2
502	tyrosyl-tRNA synthetase	630_gi|115250562	46 kDa	0	0	0	0	2
503	probable peptidase	630_gi|115251712	48 kDa	0	0	0	0	2
505	BASYS01187, 1258083-1258622 (Clockwise) Hypothetical Protein BASYS01187 putative reductase	sta_BASYS01187 115249854	21 kDa	0	0	0	0	2
507	purH, 597466-598998 (Clockwise) Hypothetical Protein purH bifunctional purine biosynthesis protein [includes: phosphoribosylaminoimidazolecarboxamide formyltransferase and IMP cyclohydrolase]	sta_BASYS00576 115249231	57 kDa	0	0	0	0	1
508	putative manganese-containing catalase	630_gi|115250607	25 kDa	0	0	0	0	1
509	putative FMN-dependent dehydrogenase	630_gi|115250297	36 kDa	0	0	0	0	2
513	phosphoglucomutase/phosphomannomutase mutase	630_gi|115249128	51 kDa	0	0	0	0	1
514	cell surface protein	630_gi|115249450	41 kDa	0	0	0	0	1
515	putative membrane protein	630_gi|115249537	42 kDa	0	0	0	0	1
516	rubredoxin oxidoreductase (desulfoferrodoxin)	630_gi|115249844	14 kDa	0	0	0	0	1
517	putative aconitase/3-isopropylmalate dehydratase	630_gi|115249850	69 kDa	0	0	0	0	1
518	ABC transporter, permease protein	630_gi|115249889	32 kDa	0	0	0	0	1
519	cold shock protein	630_gi|115249906	7 kDa	0	0	0	0	1
520	putative inorganic polyphosphate/ATP-NAD kinase	630_gi|115250070	30 kDa	0	0	0	0	1
521	putative ribonucleotide-diphosphate reductase	630_gi|115250295	83 kDa	0	0	0	0	1

522	putative bifunctional protein: peroxiredoxin/chitinase	630_gi|115250474	81 kDa	0	0	0	0	1
523	putative imidazole glycerol phosphate synthase subunit	630_gi|115250592	23 kDa	0	0	0	0	1
524	putative lipoate-protein ligase	630_gi|115250695	38 kDa	0	0	0	0	1
525	putative membrane protein	630_gi|115250814	41 kDa	0	0	0	0	1
526	biotin carboxyl carrier protein of acetyl-CoA carboxylase	630_gi|115250985	17 kDa	0	0	0	0	1
527	fructose-6-phosphate aldolase 2	630_gi|115251164	25 kDa	0	0	0	0	1
528	ribose-5-phosphate isomerase 1	630_gi|115251375	16 kDa	0	0	0	0	1
529	histidine triad nucleotide-binding protein	630_gi|115251501	13 kDa	0	0	0	0	1
530	PTS system, glucose-specific IIa component	630_gi|115251717	18 kDa	0	0	0	0	1
531	PTS system, IIc component	630_gi|115252070	39 kDa	0	0	0	0	1
532	PTS system, IIa component	630_gi|115252072	17 kDa	0	0	0	0	1
533	PTS system, IIb component	630_gi|115252122	18 kDa	0	0	0	0	1
534	phosphoglycerate kinase	630_gi|115252230	43 kDa	0	0	0	0	1
535	putative nitroreductase	630_gi|115252261	22 kDa	0	0	0	0	1
536	alanine racemase	630_gi|115252523	43 kDa	0	0	0	0	1
537	ribose-5-phosphate isomerase 2	630_gi|115252540	16 kDa	0	0	0	0	1
538	BASYS00822, 847137-847451 (Clockwise) Hypothetical Protein BASYS00822 cell surface protein (putative hemagglutinin/adhesin)	sta_BASYS00822	11 kDa	0	0	0	0	1
570	precorrin-4 C(11)-methyltransferase	630_gi|115252486	28 kDa	0	0	0	0	1
571	putative RNA-binding protein	630_gi|115250347	36 kDa	0	0	0	0	1
581	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	630_gi|115249051	17 kDa	0	0	0	0	1
582	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase	630_gi|115249069	28 kDa	0	0	0	0	1
583	putative FolD bifunctional protein includes: methylenetetrahydrofolate dehydrogenase; methenyltetrahydrofolate cyclohydrolase	630_gi|115249737	31 kDa	0	0	0	0	1
584	GntR-family transcriptional regulator	630_gi|115249901	25 kDa	0	0	0	0	1
585	conserved hypothetical protein	630_gi|115249912	57 kDa	0	0	0	0	1
586	methylglyoxal synthase	630_gi|115250185	15 kDa	0	0	0	0	1
587	branched chain amino acid transport system carrier protein	630_gi|115250294	45 kDa	0	0	0	0	1
588	cysteine desulfurase	630_gi|115250313	44 kDa	0	0	0	0	1
589	putative ATP phosphoribosyltransferase regulatory subunit	630_gi|115250588	37 kDa	0	0	0	0	1
590	histidine biosynthesis bifunctional protein includes: phosphoribosyl-AMP cyclohydrolase and phosphoribosyl-ATP pyrophosphatase	630_gi|115250595	26 kDa	0	0	0	0	1
591	probable sugar O-acetyltransferase	630_gi|115250709	20 kDa	0	0	0	0	1
592	pyrrolidone-carboxylate peptidase	630_gi|115250718	24 kDa	0	0	0	0	1
593	putative lipoprotein	630_gi|115250827	53 kDa	0	0	0	0	1
594	cytidylate kinase	630_gi|115250861	24 kDa	0	0	0	0	1
595	putative phage-related cell wall hydrolase (endolysin)	630_gi|115250945	29 kDa	0	0	0	0	1
596	putative glutaminyl-tRNA synthetase	630_gi|115251113	64 kDa	0	0	0	0	1
597	putative 2Fe-2S-binding subunit of oxidoreductase	630_gi|115251154	17 kDa	0	0	0	0	1

598	radical SAM-superfamily protein	630_gi|115251208	42 kDa	0	0	0	0	1
599	cell surface protein	630_gi|115251246	51 kDa	0	0	0	0	1
600	cinserved hypothetical protein	630_gi|115251677	47 kDa	0	0	0	0	1
601	putative histidyl-tRNA synthetase	630_gi|115251792	48 kDa	0	0	0	0	1
602	conserved hypothetical protein	630_gi|115252018	11 kDa	0	0	0	0	1
603	hypoxanthine phosphoribosyltransferase	630_gi|115252288	20 kDa	0	0	0	0	1
604	conserved hypothetical protein	630_gi|115252296	17 kDa	0	0	0	0	1
605	putative endopeptidase; pz-peptidase	630_gi|115252325	69 kDa	0	0	0	0	1
606	ATP synthase subunit gamma	630_gi|115252529	32 kDa	0	0	0	0	1
607	conserved hypothetical protein	630_gi|115252657	22 kDa	0	0	0	0	1
608	cysG, 7168-6443 (CounterClockwise) Hypothetical Protein cysG cobalt-precorrin-3b c(17)-methyltransferase	sta_BASYS00007 115252484	26 kDa	0	0	0	0	1
609	BASYS00595, 613397-613753 (Clockwise) Flagellar Protein FliS flagellar protein	sta_BASYS00595 115249244	14 kDa	0	0	0	0	1
610	BASYS00600, 618867-620744 (Clockwise) Glycosyltransferase glycosyl transferase family protein	sta_BASYS00600	72 kDa	0	0	0	0	1
611	BASYS00604, 625190-626143 (Clockwise) Hypothetical Protein BASYS00604 ornithine cyclodeaminase	sta_BASYS00604	37 kDa	0	0	0	0	1
612	udk, 898393-900057 (Clockwise) Hypothetical Protein udk putative phosphoribulokinase/uridine kinase	sta_BASYS00868 115249573	64 kDa	0	0	0	0	1
613	BASYS03517, 3762887-3762459 (CounterClockwise) Hypothetical Protein BASYS03517 putative beta-glucoside bgl operon transcription antiterminator	sta_BASYS03517 115252171	17 kDa	0	0	0	0	1
636	purine nucleoside phosphorylase	630_gi|115250257	29 kDa	0	0	0	0	1
637	putative membrane protein	630_gi|115249643	23 kDa	0	0	0	0	1
638	N-acetyl-gamma-glutamyl-phosphate reductase	630_gi|115251088	38 kDa	0	0	0	0	1
652	methenyltetrahydrofolate cyclohydrolase	630_gi|115249736	23 kDa	0	0	0	0	1
663	conserved hypothetical protein	630_gi|115252672	33 kDa	0	0	0	0	1
665	BASYS03519, 3763598-3763290 (CounterClockwise) Conserved Protein -	sta_BASYS03519 -	12 kDa	0	0	0	0	1
672	probable proton-dependent oligopeptide transporter	630_gi|115251316	50 kDa	0	0	0	0	1
673	signal peptidase I	630_gi|115249570	20 kDa	0	0	0	0	1
674	conserved hypothetical protein	630_gi|115250755	12 kDa	0	0	0	0	1
677	putative regulatory protein	630_gi|115250392	18 kDa	0	0	0	0	1
680	conserved hypothetical protein	630_gi|115252343	59 kDa	0	0	0	0	1
682	conserved hypothetical protein	630_gi|115249648	42 kDa	0	0	0	0	1
691	BASYS02943, 3088650-3087466 (CounterClockwise) Acetyltransferase GNAT Family -	sta_BASYS02943 115251530	46 kDa	0	0	0	0	1
693	abgT, 2475250-2476803 (Clockwise) Hypothetical Protein abgT aminobenzoyl-glutamate transport protein	sta_BASYS02378 -	55 kDa	0	0	0	0	1
694	putative component of D-ornithine aminomutase	630_gi|115249458	52 kDa	0	0	0	0	1
695	toxin B	630_gi|115249675	270 kDa	0	0	0	0	1

699	BASYS03375, 3601708-3602172 (Clockwise) Hypothetical Protein BASYS03375 hypothetical protein	sta_BASYS03375 115251932	18 kDa	0	0	0	0	1
710	peptide deformylase 2	630_gi|115251641	16 kDa	0	0	0	0	1
713	putative RNA-binding protein	630_gi|115249154	81 kDa	0	0	0	0	1

Table A6 – *Proteins identified In strain B-1 ProteoMiner treated broth extracts by 1D GE followed by LC-MS/MS analysis with identification number, predicted mass and number of unique peptides detected in each sample. The number of proteins of different types are show as follows; metabolic (m), surface and virulence (s/v), hypothetical (h), regulatory (R), transporter (t) and ribosomal (r).*

	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	165 protiens m=107, r=32, h=8, R=13, s/v=4, t=1			1	2	1	2	1
	All samples							
	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405	41 kDa	23	18	15	14	19
2	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249404	42 kDa	20	19	14	14	22
3	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	630_gi|115249401	44 kDa	18	16	6	6	16
4	pyruvate-flavodoxin oxidoreductase	630_gi|115251733	129 kDa	19	17	11	14	17
5	30S ribosomal protein S7	630_gi|115249073	18 kDa	13	11	18	13	5
6	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249403	47 kDa	20	16	9	12	16
7	electron transfer flavoprotein alpha-subunit	630_gi|115249407	37 kDa	10	6	5	5	17
8	50S ribosomal protein L19	630_gi|115250291	13 kDa	5	4	8	7	1
9	electron transfer flavoprotein beta-subunit	630_gi|115249406	29 kDa	9	7	6	5	9
10	50S ribosomal protein L15	630_gi|115249096	16 kDa	3	3	8	5	2
11	50S ribosomal protein L5	630_gi|115249089	20 kDa	12	7	14	16	9
12	putative amino acid aminotransferase	630_gi|115252729	45 kDa	13	13	5	6	16
13	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184	68 kDa	14	12	8	11	17
14	putative ruberythrin	630_gi|115250515	20 kDa	5	4	3	3	6
15	adhE, 3383815-3381161 (CounterClockwise) Hypothetical Protein adhE aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase	stb_BASYS03189 115252023	97 kDa	17	16	8	9	19
16	30S ribosomal protein S4	630_gi|115249105	24 kDa	9	9	12	11	7
17	butyryl-CoA dehydrogenase	630_gi|115250075	41 kDa	14	13	7	10	10
18	putative oxidoreductase, thiamine diP-binding subunit	630_gi|115249125	39 kDa	6	7	7	5	6
19	glyceraldehyde-3-phosphate dehydrogenase 2	630_gi|115252231	36 kDa	9	8	8	8	10
20	50S ribosomal protein L1	630_gi|115249066	25 kDa	7	6	4	6	4
21	succinate-semialdehyde dehydrogenase NAD(P)+	630_gi|115251397	51 kDa	9	9	7	10	21
22	30S ribosomal protein S9	630_gi|115249113	15 kDa	7	5	5	6	4
23	50S ribosomal protein L24	630_gi|115249088	11 kDa	2	2	4	4	2
24	NAD-specific glutamate dehydrogenase	630_gi|115249189	46 kDa	10	10	3	5	16
25	glycine/sarcosine/betaine reductase complex component C beta subunit	630_gi|115251404	55 kDa	6	6	5	5	8
26	30S ribosomal protein S5	630_gi|115249094	18 kDa	6	6	10	8	2
27	50S ribosomal protein L6	630_gi|115249092	20 kDa	5	3	10	13	3
28	acetyl-CoA acetyltransferase	630_gi|115250080	41 kDa	6	6	7	8	12
29	30S ribosomal protein S10	630_gi|115249076	12 kDa	4	4	4	4	3

30	glycine/sarcosine/betaine reductase complex component C alpha subunit	630_gi|115251403	41 kDa	10	9	5	6	11
31	eutG, 2931498-2930332 (CounterClockwise) Hypothetical Protein eutG -	sta_BASYS02789 -	44 kDa	8	8	4	4	8
32	50S ribosomal protein L13	630_gi|115249112	16 kDa	1	1	4	4	1
33	putative translation elongation factor	630_gi|115249025	72 kDa	12	12	6	10	11
34	50S ribosomal protein L4	630_gi|115249078	24 kDa	5	4	5	6	3
35	60 kDa chaperonin	630_gi|115249204	58 kDa	15	16	6	6	20
36	elongation factor TU	630_gi|115249061	44 kDa	9	6	3	4	9
37	rubrerythrin	630_gi|115249842	21 kDa	5	4	2	3	4
38	aspartate aminotransferase	630_gi|115250375	45 kDa	11	13	5	7	11
39	50S ribosomal protein L3	630_gi|115249077	22 kDa	6	5	4	5	5
40	yqjB, 3369095-3367134 (CounterClockwise) Uncharacterized protein yqjB hypothetical protein	sta_BASYS03185 115252019	72 kDa	6	4	11	10	1
41	(R)-2-hydroxyisocaproate dehydrogenase	630_gi|115249400	37 kDa	5	4	3	5	9
42	DNA-directed RNA polymerase beta chain	630_gi|115249070	139 kDa	16	12	5	10	4
44	50S ribosomal protein L21	630_gi|115250193	11 kDa	4	5	4	4	3
45	norV, 2106146-2108677 (Clockwise) Hypothetical Protein norV putative nitric oxide reductase flavoprotein	sta_BASYS02024 115250664	94 kDa	11	10	0	1	13
46	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC flagellin subunit	sta_BASYS00598 115249247	34 kDa	8	6	8	6	9
229	flagellin subunit	630_gi|115249247	31 kDa	1	1	1	1	0
47	30S ribosomal protein S13	630_gi|115249103	14 kDa	4	4	5	5	2
48	putative subunit of oxidoreductase	630_gi|115249127	20 kDa	5	3	1	1	3
49	30S ribosomal protein S20	630_gi|115251527	10 kDa	2	2	2	2	1
50	glycine reductase complex component B alpha and beta subunits	630_gi|115251407	46 kDa	9	5	3	3	12
51	30S ribosomal protein S11	630_gi|115249104	14 kDa	4	3	5	3	2
52	stage 0 sporulation protein A	630_gi|115250247	31 kDa	4	2	9	7	5
53	threonine dehydratase catabolic	630_gi|115251567	43 kDa	4	4	6	6	9
54	activator of 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249402	29 kDa	4	4	2	3	6
55	30S ribosomal protein S8	630_gi|115249091	15 kDa	5	2	5	4	3
56	putative anaerobic nitric oxide reductase flavorubredoxin	630_gi|115250189	45 kDa	6	5	1	1	3
57	putative rubrerythrin	630_gi|115250565	20 kDa	3	2	1	1	2
58	50S ribosomal protein L2	630_gi|115249080	30 kDa	5	4	4	4	2
59	yhaM, 3836692-3835367 (CounterClockwise) Hypothetical Protein yhaM conserved hypothetical protein	stb_BASYS03586 115252289	47 kDa	5	4	4	5	10
60	30S ribosomal protein S3	630_gi|115249083	30 kDa	2	3	4	3	1
61	toxin A	630_gi|115249677	308 kDa	8	2	2	1	22
62	ATP synthase beta chain	630_gi|115252528	50 kDa	5	4	2	3	4
63	50S ribosomal protein L22	630_gi|115249082	12 kDa	5	4	7	7	2
64	phosphate butyryltransferase	630_gi|115249121	32 kDa	7	4	4	2	5
65	DNA-directed RNA polymerase beta' chain	630_gi|115249071	130 kDa	8	5	4	7	1
66	putative aminoacyl-histidine dipeptidase	630_gi|115249724	53 kDa	7	5	0	1	7
67	glyS, 3040737-3038671 (CounterClockwise) Hypothetical Protein glyS glycyl-tRNA synthetase beta chain	sta_BASYS02895 115251485	78 kDa	9	8	6	9	4

69	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,methyltransferase subunit	630_gi|115249744	29 kDa	7	7	1	1	3
70	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, alpha subunit	630_gi|115249743	50 kDa	5	4	2	3	6
71	translation initiation factor IF-3	630_gi|115249701	19 kDa	1	1	6	7	1
72	ATP synthase alpha chain	630_gi|115252530	55 kDa	5	6	1	0	8
73	gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase	630_gi|115251396	55 kDa	3	2	0	1	11
75	putative proline racemase	630_gi|115252294	36 kDa	6	4	2	4	7
76	4-hydroxybutyrate CoA transferase	630_gi|115251394	48 kDa	3	2	1	2	11
77	translation elongation factor G	630_gi|115249074	76 kDa	7	3	0	3	13
78	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079	31 kDa	5	4	3	3	4
79	glycine reductase complex component B gamma subunit	630_gi|115251405	47 kDa	1	2	2	2	7
80	thioredoxin reductase	630_gi|115251409	34 kDa	6	3	1	1	10
81	30S ribosomal protein S19	630_gi|115249081	11 kDa	1	1	2	2	1
82	formate--tetrahydrofolate ligase	630_gi|115249735	60 kDa	2	3	3	4	9
83	butyrate kinase	630_gi|115249122	39 kDa	3	2	3	3	13
85	putative glutamate synthase NADPH small chain	630_gi|115250578	50 kDa	5	3	1	1	7
86	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	630_gi|115249053	55 kDa	3	3	2	2	3
87	putative ornithine cyclodeaminase	630_gi|115249560	36 kDa	2	3	5	3	2
88	putative oxidoreductase, NAD/FAD binding subunit	630_gi|115249186	46 kDa	5	5	2	2	10
89	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	630_gi|115249740	49 kDa	2	2	5	4	2
94	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872	59 kDa	4	4	2	2	6
95	chaperone protein	630_gi|115251515	66 kDa	1	1	2	3	9
96	pyruvate, phosphate dikinase	630_gi|115251462	97 kDa	8	5	0	1	2
97	leucyl-tRNA synthetase	630_gi|115251575	92 kDa	9	7	0	3	4
98	putative dehydrogenase, electron transfer subunit	630_gi|115250577	33 kDa	4	4	0	1	4
100	DNA-directed RNA polymerase alpha chain	630_gi|115249106	35 kDa	5	5	4	2	6
102	30S ribosomal protein S2	630_gi|115251194	27 kDa	4	3	3	4	2
103	ribose-phosphate pyrophosphokinase	630_gi|115252575	34 kDa	4	2	2	1	2
104	enolase	630_gi|115252227	46 kDa	4	4	2	2	7
105	trigger factor	630_gi|115252362	48 kDa	2	1	3	4	7
106	electron transfer flavoprotein alpha-subunit	630_gi|115250077	36 kDa	5	3	1	0	5
107	formate acetyltransferase	630_gi|115249776	84 kDa	6	3	1	0	8
108	putative propanediol utilization protein	630_gi|115251734	20 kDa	3	1	4	4	5
109	putative 2-hydroxyacyl-CoA dehydratase	630_gi|115250793	44 kDa	3	2	1	2	3
110	electron transfer flavoprotein beta-subunit	630_gi|115250076	29 kDa	4	4	0	1	4
111	30S ribosomal protein S17	630_gi|115249086	10 kDa	2	2	2	3	1
112	ferredoxin	630_gi|115249124	8 kDa	3	3	2	3	2
113	50S ribosomal protein L10	630_gi|115249067	19 kDa	5	5	1	3	3
115	putative phosphate butyryltransferase	630_gi|115249732	33 kDa	4	3	3	2	4

120	GTP-sensing transcriptional pleiotropic repressor	630_gi|115250309	29 kDa	2	1	4	3	3
123	putative electron transfer protein	630_gi|115252303	47 kDa	2	1	5	6	2
124	30S ribosomal protein S6	630_gi|115252728	11 kDa	3	2	3	3	3
125	50S ribosomal protein L9	630_gi|115252722	17 kDa	1	1	3	6	1
126	ucpA, 407353-408141 (Clockwise) Hypothetical Protein ucpA NADP-dependent 7-alpha-hydroxysteroid dehydrogenase	sta_BASYS00404 115249069	28 kDa	4	3	0	1	5
129	putative transcriptional regulator	630_gi|115250535	23 kDa	2	2	2	2	2
130	glucosamine--fructose-6-phosphate aminotransferase isomerizing	630_gi|115249129	67 kDa	3	3	0	1	7
134	metallo beta-lactamase superfamily protein	630_gi|115250323	62 kDa	4	4	2	3	3
135	conserved hypothetical protein	630_gi|115249288	15 kDa	2	2	3	3	2
139	putative universal stress protein	630_gi|115249829	15 kDa	5	3	1	1	2
141	UDP-glucose 4-epimerase	630_gi|115251765	37 kDa	3	2	4	1	4
142	50S ribosomal protein L29	630_gi|115249085	8 kDa	2	2	2	2	1
143	valS, 4034578-4031912 (CounterClockwise) Valyl-tRNA synthetase valyl-tRNA synthetase	sta_BASYS03755 115252312	103 kDa	3	2	1	1	1
144	methionyl-tRNA synthetase	630_gi|115252601	74 kDa	3	3	0	1	3
145	putative RNA-binding protein	630_gi|115249151	21 kDa	1	1	1	1	1
146	putative nogalamycin resistance protein	630_gi|115252273	83 kDa	1	1	1	1	1
147	ATP synthase subunit delta	630_gi|115252531	21 kDa	2	0	3	4	2
148	agaY, 769977-769051 (CounterClockwise) Hypothetical Protein agaY putative fructose-bisphosphate aldolase	sta_BASYS00749 115249409	33 kDa	3	3	0	1	5
150	tellurium resistance protein	630_gi|115250846	21 kDa	1	0	3	6	1
153	pyruvate kinase	630_gi|115252454	63 kDa	2	2	1	1	2
158	BASYS04047, 4321907-4321278 (CounterClockwise) Hypothetical Protein BASYS04047 SpoIIIJ-associated protein	stb_BASYS04047 115252741	24 kDa	3	2	1	2	2
159	conserved hypothetical protein	630_gi|115252075	17 kDa	2	2	2	2	1
160	putative FMN-dependent dehydrogenase	630_gi|115250297	36 kDa	2	2	2	0	3
161	putative NUDIX-family hydrolase	630_gi|115249807	17 kDa	3	0	2	2	2
162	tyrosyl-tRNA synthetase	630_gi|115250562	46 kDa	1	2	0	3	5
163	hypothetical protein	630_gi|115249824	14 kDa	1	1	1	1	1
164	single-strand binding protein	630_gi|115252727	16 kDa	4	3	0	1	1
167	glyceraldehyde-3-phosphate dehydrogenase 1	630_gi|115250813	37 kDa	1	2	1	2	2
168	cysteinyl-tRNA synthetase	630_gi|115249055	54 kDa	1	1	1	2	4
169	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, small subunit	630_gi|115249742	34 kDa	1	1	1	1	1
170	isoleucyl-tRNA synthetase	630_gi|115251669	120 kDa	4	1	0	1	1
171	50S ribosomal protein L7/L12	630_gi|115249068	13 kDa	2	2	1	3	1
175	nitroreductase-family protein	630_gi|115250155	33 kDa	1	1	0	1	2
176	elongation factor Ts	630_gi|115251193	33 kDa	1	0	0	1	7
179	putative tellurite resistance protein	630_gi|115250680	44 kDa	2	2	1	2	3
180	putative 30S ribosomal protein S1	630_gi|115250007	48 kDa	1	0	2	2	3
183	putative aminopeptidase	630_gi|115250096	51 kDa	1	1	0	1	1
186	ribosome recycling factor	630_gi|115251191	21 kDa	1	1	1	1	2
187	butyrate kinase	630_gi|115251431	39 kDa	1	1	1	0	2

193	DNA gyrase subunit A	630_gi|115249009	91 kDa	2	1	2	2	2
196	DNA polymerase III, beta chain	630_gi|115249005	46 kDa	0	2	1	1	2
201	acetate kinase	630_gi|115250207	43 kDa	1	0	0	1	8
204	putative cold shock protein	630_gi|115250391	7 kDa	1	1	1	1	1
205	ATP-dependent Clp protease ATP-binding subunit	630_gi|115252360	46 kDa	2	1	0	1	2
212	DNA-binding protein HU	630_gi|115252557	10 kDa	1	1	1	2	1
214	electron transfer flavoprotein beta-subunit	630_gi|115249821	29 kDa	1	1	1	1	1
215	uridylate kinase	630_gi|115251192	26 kDa	1	1	0	1	2
220	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	630_gi|115251182	38 kDa	1	1	1	1	1
221	putative dinitrogenase iron-molybdenum cofactor	630_gi|115250736	13 kDa	3	1	0	1	1
223	putative selenocysteine lyase	630_gi|115252735	42 kDa	1	1	0	1	1
227	seryl-tRNA synthetase	630_gi|115249017	49 kDa	1	1	1	1	1
228	orotate phosphoribosyltransferase	630_gi|115249197	21 kDa	1	1	0	1	2
231	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	630_gi|115249132	45 kDa	1	1	0	1	1
232	aspartate aminotransferase	630_gi|115249115	44 kDa	1	1	0	1	1
246	conserved hypothetical protein	630_gi|115250551	24 kDa	0	1	1	2	1
247	glucose-6-phosphate isomerase	630_gi|115252341	51 kDa	0	1	1	1	2
266	threonine synthase	630_gi|115251172	56 kDa	1	1	0	1	1
273	10 kDa chaperonin	630_gi|115249203	10 kDa	1	0	0	1	1
287	uracil phosphoribosyltransferase	630_gi|115252539	23 kDa	1	0	1	1	1
296	conserved hypothetical protein	630_gi|115250191	27 kDa	0	1	0	1	1
297	nitrilase (carbon-nitrogen hydrolase)	630_gi|115251895	34 kDa	2	0	1	0	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	92 protiens m=68, r=0, h=10, R=9, s/v=1, t=4			1	2	1	2	1
	Ebeads + crude							
101	peptidase	630_gi|115251315	68 kDa	7	5	0	0	6
114	GMP synthase glutamine-hydrolyzing	630_gi|115249206	57 kDa	3	4	0	0	6
116	indolepyruvate oxidoreductase subunit	630_gi|115251433	66 kDa	6	4	0	0	2
117	proline reductase subunit proprotein	630_gi|115252300	68 kDa	4	1	0	0	6
119	ABC transporter, substrate-binding lipoprotein	630_gi|115249887	36 kDa	4	2	0	0	7
127	succinyl-CoA:coenzyme A transferase	630_gi|115251398	56 kDa	4	3	0	0	7
131	threonyl-tRNA synthetase	630_gi|115249589	74 kDa	3	3	0	0	2
132	phenylalanyl-tRNA synthetase beta chain	630_gi|115249716	89 kDa	4	3	0	0	5
140	AsnC-family transcriptional regulator	630_gi|115252605	16 kDa	5	3	0	0	3
151	putative glycosyl hydrolase	630_gi|115252069	100 kDa	4	4	0	0	4
152	chaperone protein (heat shock protein)	630_gi|115249282	75 kDa	3	2	0	0	6
154	V-type sodium ATP synthase subunit A	630_gi|115252012	66 kDa	3	3	0	0	3
155	phosphoenolpyruvate-protein phosphotransferase	630_gi|115251808	63 kDa	3	2	0	0	5
157	glycogen synthase	630_gi|115249898	56 kDa	4	4	0	0	1
165	putative peptidase	630_gi|115252582	40 kDa	2	1	0	0	5
174	cell division protein	630_gi|115251697	41 kDa	2	2	0	0	3
177	putative transaldolase	630_gi|115251384	23 kDa	2	0	0	0	5
181	conserved hypothetical protein	630_gi|115252618	28 kDa	3	3	0	0	2
182	proline iminopeptidase	630_gi|115252096	34 kDa	2	2	0	0	2
184	aspartate--ammonia ligase	630_gi|115251746	39 kDa	3	2	0	0	1

188	manganese-dependent inorganic pyrophosphatase	630_gi|115249342	59 kDa	4	3	0	0	2
189	GntR-family transcriptional regulator	630_gi|115252628	25 kDa	3	3	0	0	2
190	putative aliphatic sulfonates ABC transporter,substrate-binding lipoprotein	630_gi|115250525	36 kDa	2	2	0	0	3
191	putative amidohydrolase	630_gi|115252278	42 kDa	2	2	0	0	3
194	rubredoxin oxidoreductase (desulfoferrodoxin)	630_gi|115249844	14 kDa	2	2	0	0	3
195	conserved hypothetical protein	630_gi|115249797	34 kDa	1	2	0	0	2
197	oligopeptide ABC transporter, substrate-binding protein	630_gi|115251723	58 kDa	1	1	0	0	4
199	cyclopropane-fatty-acyl-phospholipid synthase	630_gi|115249187	46 kDa	1	1	0	0	2
200	glycyl-tRNA synthetase alpha chain	630_gi|115251486	34 kDa	3	1	0	0	1
206	NifU-like protein	630_gi|115250314	16 kDa	2	1	0	0	2
208	putative 1-(5-phosphoribosyl)-5-(5-phosphoribosylamino)methylidene amino imidazole-4-carboxamide isomerase	630_gi|115250593	27 kDa	2	2	0	0	2
210	dihydrodipicolinate synthase	630_gi|115252282	32 kDa	2	0	0	0	4
216	6-phosphofructokinase	630_gi|115252455	34 kDa	1	1	0	0	4
219	phosphoglycerate kinase	630_gi|115252230	43 kDa	1	0	0	0	5
222	LacI-family transcriptional regulator (catabolite control protein)	630_gi|115250088	38 kDa	1	2	0	0	2
224	putative alanyl-tRNA synthetase	630_gi|115250316	98 kDa	3	1	0	0	1
225	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	630_gi|115249733	69 kDa	1	1	0	0	1
226	heat shock protein	630_gi|115251516	24 kDa	1	0	0	0	4
233	BASYS00477, 430146-431189 (Clockwise) Hypothetical Protein BASYS00477 -	stb_BASYS00477 -	41 kDa	2	1	0	0	1
234	putative glutamine amidotransferase	630_gi|115252147	28 kDa	2	0	0	0	2
235	putative DNA-binding protein	630_gi|115249180	23 kDa	2	0	0	0	2
236	ribulose-phosphate 3-epimerase	630_gi|115251631	24 kDa	2	0	0	0	2
239	putative sulfonate ABC transporter,solute-binding lipoprotein	630_gi|115251417	40 kDa	2	1	0	0	0
240	transketolase	630_gi|115251376	32 kDa	1	0	0	0	4
241	conserved hypothetical protein	630_gi|115251115	24 kDa	1	0	0	0	3
248	thioredoxin	630_gi|115250733	12 kDa	2	1	0	0	2
249	anti-sigma F factor	630_gi|115249788	16 kDa	2	2	0	0	1
252	putative phosphatase	630_gi|115251616	60 kDa	1	1	0	0	2
253	putative acetyltransferase	630_gi|115249853	19 kDa	2	1	0	0	1
255	anti-sigma F factor antagonist	630_gi|115249787	13 kDa	1	1	0	0	1
256	putative amidohydrolase	630_gi|115249796	41 kDa	1	1	0	0	1
258	probable peptidase	630_gi|115251664	40 kDa	2	0	0	0	3
260	tellurium resistance protein	630_gi|115250677	23 kDa	2	0	0	0	2
261	putative alanine racemase	630_gi|115251674	27 kDa	1	0	0	0	4
262	tellurium resistance protein	630_gi|115250676	22 kDa	1	0	0	0	3
267	glycerol dehydratase	630_gi|115250149	89 kDa	1	2	0	0	2
269	histidine biosynthesis bifunctional protein includes: phosphoribosyl-AMP cyclohydrolase and phosphoribosyl-ATP pyrophosphatase	630_gi|115250595	26 kDa	1	1	0	0	4
270	conserved hypothetical protein	630_gi|115251019	48 kDa	1	1	0	0	2
271	putative molybdopterin biosynthesis protein	630_gi|115251123	37 kDa	1	1	0	0	2

272	putative aldose epimerase	630_gi|115252500	34 kDa	1	1	0	0	2
274	low-specificity L-threonine aldolase	630_gi|115251648	38 kDa	1	0	0	0	2
275	conserved hypothetical protein	630_gi|115249812	26 kDa	0	1	0	0	2
276	cysteine desulfurase	630_gi|115250313	44 kDa	1	0	0	0	3
278	BASYS02513, 2601517-2601618 (Clockwise) Hypothetical Protein BASYS02513 putative membrane protein	stb_BASYS02513 115251345	4 kDa	1	1	0	0	1
279	putative indolepyruvate oxidoreductase subunit	630_gi|115251432	21 kDa	1	0	0	0	1
288	glutamyl-tRNA synthetase	630_gi|115249054	57 kDa	1	1	0	0	2
290	putative ATP-binding protein	630_gi|115250774	29 kDa	1	1	0	0	1
291	BASYS00737, 756750-756154 (CounterClockwise) Deoxyribonuclease TatD Family Radical SAM-superfamily protein	sta_BASYS00737 115249398	23 kDa	1	1	0	0	1
292	BASYS03545, 3789976-3788798 (CounterClockwise) Hypothetical Protein BASYS03545 -	stb_BASYS03545 -	46 kDa	1	1	0	0	1
294	L-serine dehydratase	630_gi|115252279	43 kDa	1	1	0	0	1
298	BASYS02738, 2836977-2838605 (Clockwise) Hypothetical Protein BASYS02738 L-aspartate-beta-decarboxylase	stb_BASYS02738 115251568	62 kDa	1	0	0	0	2
299	putative ATP-binding protein	630_gi|115250457	25 kDa	1	0	0	0	2
302	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	630_gi|115252499	37 kDa	1	0	0	0	1
305	cell-division initiation protein	630_gi|115251670	20 kDa	1	0	0	0	2
313	ferritin	630_gi|115251248	20 kDa	1	1	0	0	1
314	putative aspartyl aminopeptidase	630_gi|115249338	48 kDa	1	1	0	0	1
324	putative aspartate aminotransferase	630_gi|115251434	47 kDa	0	1	0	0	1
325	putative 5-nitroimidazole reductase	630_gi|115250500	18 kDa	1	0	0	0	1
342	rRNA adenine N-6-methyltransferase (erythromycin resistance protein)	630_gi|115251058	29 kDa	1	1	0	0	1
343	conserved hypothetical protein	630_gi|115251617	13 kDa	1	1	0	0	1
345	utp--glucose-1-phosphate uridylyltransferase (general stress protein 33)	630_gi|115252549	36 kDa	2	0	0	0	1
347	conserved hypothetical protein	630_gi|115252038	35 kDa	1	0	0	0	1
351	putative DNA-binding protein	630_gi|115249031	40 kDa	0	1	0	0	1
355	probable sugar O-acetyltransferase	630_gi|115250709	20 kDa	1	0	0	0	1
356	probable amino acid racemase	630_gi|115251082	26 kDa	1	0	0	0	1
357	delta-aminolevulinic acid dehydratase	630_gi|115252479	36 kDa	1	0	0	0	1
358	bifunctional adenosylcobalamin biosynthesis protein includes: adenosylcobinamide kinase and adenosylcobinamide-phosphate guanylyltransferase	630_gi|115252498	22 kDa	1	0	0	0	1
360	phenylalanyl-tRNA synthetase alpha chain	630_gi|115249715	38 kDa	1	0	0	0	1
362	putative metal dependent phosphohydrolase	630_gi|115251577	22 kDa	1	0	0	0	1
392	RecA protein (recombinase A)	630_gi|115250364	37 kDa	1	0	0	0	1
397	conserved hypothetical protein	630_gi|115252731	10 kDa	1	0	0	0	1
398	putative aspartyl-tRNA synthetase	630_gi|115251791	67 kDa	1	0	0	0	1

	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	60 protiens m=37, r=1, h=16, R=6, s/v=0, t=0			1	2	1	2	1
	Ebeads only							
268	BASYS00003, 1394-2224 (Clockwise) Hypothetical Protein BASYS00003	Stbphage1_BASYS00003	33 kDa	2	2	0	0	
166	ribonuclease R	630_gi|115252220	82 kDa	5	5	0	0	0
202	conserved hypothetical protein	630_gi|115250197	14 kDa	3	2	0	0	0
209	putative GTP-binding protein	630_gi|115251683	50 kDa	3	3	0	0	0
259	putative glycine cleavage system H protein	630_gi|115249746	14 kDa	2	2	0	0	0
264	4-aminobutyrate aminotransferase	630_gi|115251211	48 kDa	1	1	0	0	0
295	putative transcriptional repressor	630_gi|115252578	32 kDa	2	1	0	0	0
315	LysR-family transcriptional regulator	630_gi|115250103	35 kDa	2	1	0	0	0
318	conserved hypothetical protein	630_gi|115251507	39 kDa	1	1	0	0	0
319	BASYS01206, 1239342-1239710 (Clockwise) Hypothetical Protein BASYS01206	stb_BASYS01206 -	14 kDa	1	1	0	0	0
322	SsrA-binding protein	630_gi|115252213	19 kDa	1	1	0	0	0
340	conserved hypothetical protein	630_gi|115252645	33 kDa	1	1	0	0	0
344	pyruvate carboxylase	630_gi|115249024	129 kDa	2	1	0	0	0
348	formylglycinamidine ribonucleotide synthetase	630_gi|115249233	141 kDa	1	0	0	0	0
349	hypothetical protein	630_gi|115249752	14 kDa	1	1	0	0	0
350	putative aromatic amino acid aminotransferase	630_gi|115251881	45 kDa	1	1	0	0	0
379	translation initiation factor IF-2	630_gi|115250345	70 kDa	1	0	0	0	0
384	putative RNA-binding mediating protein	630_gi|115249006	8 kDa	1	1	0	0	0
385	metallo-beta-lactamase superfamily protein	630_gi|115249353	35 kDa	1	1	0	0	0
386	conserved hypothetical protein	630_gi|115251511	30 kDa	1	1	0	0	0
387	putative amino acid racemase	630_gi|115251872	40 kDa	1	1	0	0	0
388	putative glycosyl hydrolase	630_gi|115252139	101 kDa	1	1	0	0	0
389	BASYS00184, 156070-156327 (Clockwise) Hypothetical Protein BASYS00184	stb_BASYS00184 -	10 kDa	1	1	0	0	0
399	BASYS00600, 618867-620744 (Clockwise) Glycosyltransferase	sta_BASYS00600	72 kDa	1	1	0	0	0
400	rod shape-determining protein	630_gi|115250177	38 kDa	1	1	0	0	0
409	putative iron-sulfur protein	630_gi|115249747	72 kDa	2	0	0	0	0
414	putative fructose-1,6-bisphosphatase	630_gi|115250224	76 kDa	0	1	0	0	0
415	putative ribonucleotide-diphosphate reductase	630_gi|115250295	83 kDa	0	1	0	0	0
416	septum site-determining protein (cell division inhibitor)	630_gi|115250182	29 kDa	1	0	0	0	0
446	S-ribosylhomocysteinase	630_gi|115252662	17 kDa	1	0	0	0	0
449	conserved hypothetical protein	630_gi|115250029	32 kDa	0	1	0	0	0
450	putative pyridoxine kinase	630_gi|115250389	31 kDa	0	1	0	0	0
451	N-carbamoyl-L-amino acid hydrolase	630_gi|115251081	44 kDa	0	1	0	0	0
452	putative carbonic anhydrase	630_gi|115251267	24 kDa	0	1	0	0	0
453	D-alanine--poly(phosphoribitol) ligase subunit 2 (D-alanyl carrier protein)	630_gi|115251904	9 kDa	0	1	0	0	0
454	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252652	119 kDa	0	1	0	0	0
455	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252654	119 kDa	0	1	0	0	0
456	aminopeptidase	630_gi|115252659	46 kDa	0	1	0	0	0

457	BASYS03059, 3222482-3221574 (CounterClockwise) Hypothetical Protein NE	stb_BASYS03059 -	36 kDa	0	1	0	0	0
458	conserved hypothetical protein	630_gi|115249592)	22 kDa	1	0	0	0	0
459	methylglyoxal synthase	630_gi|115250185	15 kDa	1	0	0	0	0
460	GTP-binding protein	630_gi|115251490	34 kDa	1	0	0	0	0
461	conserved hypothetical protein	630_gi|115251612	20 kDa	1	0	0	0	0
462	putative GTP pyrophosphokinase	630_gi|115251796	85 kDa	1	0	0	0	0
463	proline reductase	630_gi|115252298	26 kDa	1	0	0	0	0
464	ATP-dependent protease La	630_gi|115252357	90 kDa	1	0	0	0	0
465	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252465	18 kDa	1	0	0	0	0
466	putative nicotinate phosphribosyltransferase	630_gi|115252604	54 kDa	1	0	0	0	0
467	30S ribosomal protein S18	630_gi|115252726	9 kDa	1	0	0	0	0
468	rmlA1, 3228225-3227347 (CounterClockwise) Hypothetical Protein rmlA1	stb_BASYS03065 -	33 kDa	1	0	0	0	0
470	xanthine phosphoribosyltransferase	630_gi|115251385	21 kDa	1	0	0	0	0
508	conserved hypothetical protein	630_gi|115249912	57 kDa	1	0	0	0	0
510	conserved hypothetical protein	630_gi|115250233	13 kDa	1	0	0	0	0
515	3-hydroxybutyryl-CoA dehydratase	630_gi|115250078	28 kDa	1	0	0	0	0
516	putative transcription antiterminator	630_gi|115251388	82 kDa	1	0	0	0	0
517	putative exported carboxy-terminal processing protease	630_gi|115252515	42 kDa	0	1	0	0	0
518	N-acetyl-gamma-glutamyl-phosphate reductase	630_gi|115251088	38 kDa	1	0	0	0	0
520	conserved hypothetical protein	630_gi|115252126	20 kDa	1	0	0	0	0
539	BASYS00604, 625190-626143 (Clockwise) Hypothetical Protein BASYS00604 ornithine cyclodeaminase	sta_BASYS00604 -	37 kDa	1	0	0	0	0
540	putative glycosyl hydrolase	630_gi|115251625	103 kDa	1	0	0	0	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	28 protiens m=10, r=10, h=4, R=4, s/v=0, t=0			1	2	1	2	1
	Ebeads and SE beads							
84	50S ribosomal protein L27	630_gi|115250195	10 kDa	3	1	3	5	0
90	30S ribosomal protein S15	630_gi|115250352	10 kDa	5	4	6	6	0
91	putative ribosomal protein	630_gi|115249100	11 kDa	3	2	4	6	0
92	50S ribosomal protein L16	630_gi|115249084	16 kDa	1	2	2	2	0
93	30S ribosomal protein S16	630_gi|115250287	10 kDa	1	1	5	4	0
99	50S ribosomal protein L17	630_gi|115249107	13 kDa	1	1	2	2	0
121	50S ribosomal protein L18	630_gi|115249093	14 kDa	1	1	3	4	0
122	50S ribosomal protein L11	630_gi|115249065	15 kDa	1	1	1	1	0
133	50S ribosomal protein L14	630_gi|115249087	13 kDa	0	1	3	4	0
136	PTS system, IIb component	630_gi|115252071	11 kDa	1	1	2	3	0
149	50S ribosomal protein L20	630_gi|115249703	13 kDa	1	1	1	2	0
172	BASYS01737, 1767940-1768269 (Clockwise) Hypothetical Protein BASYS01737 -	stb_BASYS01737 -	13 kDa	2	2	2	3	0
173	transcription termination factor Rho	630_gi|115252548	61 kDa	3	3	1	1	0
192	putative serine hydroxymethyltransferase	630_gi|115251778	46 kDa	1	1	1	1	0
198	GntR-family transcriptional regulator	630_gi|115250423	24 kDa	1	0	2	2	0

207	3-oxoacyl-acyl-carrier-protein synthase II	630_gi|115250217	44 kDa	2	1	1	1	0
213	conserved hypothetical protein	630_gi|115251576	13 kDa	2	1	1	2	0
217	putative beta-lactamase repressor	630_gi|115251442	15 kDa	1	0	2	1	0
230	conserved hypothetical protein	630_gi|115250755	12 kDa	1	1	1	1	0
238	geranyltranstransferase	630_gi|115250238	33 kDa	0	1	1	1	0
251	phosphoribosylaminoimidazole carboxylase catalytic subunit	630_gi|115249226	17 kDa	1	1	0	2	0
254	adenylosuccinate lyase	630_gi|115250370	55 kDa	1	1	0	1	0
289	ATP-dependent Clp protease proteolytic subunit	630_gi|115252361	21 kDa	1	1	0	1	0
326	BASYS02591, 2703866-2703633 (CounterClockwise) Hypothetical Protein BASYS02591 putative regulatory protein	sta_BASYS02591 115251185	9 kDa	0	1	1	0	0
346	conserved hypothetical protein	630_gi|115250624	11 kDa	1	0	1	0	0
390	putative component of D-ornithine aminomutase	630_gi|115249458	52 kDa	1	0	1	0	0
391	putative D-tyrosyl-tRNA protein	630_gi|115251795	16 kDa	1	0	0	1	0
396	preprotein translocase SecA subunit	630_gi|115249152	102 kDa	1	0	0	1	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	7 protiens m=1, r=0, h=1, R=0, s/v=5, t=0			1	2	1	2	1
	SE beads and crude							
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
				1	2	1	2	1
74	ydhO, 1281410-1282705 (Clockwise) Hypothetical Protein ydhO putative cell wall hydrolase	stb_BASYS01250 115250166	45 kDa	0	0	10	6	3
257	BASYS03070, 3234507-3232624 (CounterClockwise) Hypothetical Protein BASYS03070 cell surface protein	stb_BASYS03070 115251842	68 kDa	0	0	1	1	1
277	conserved hypothetical protein	630_gi|115251464	23 kDa	0	0	1	0	2
293	putative signal recognition particle	630_gi|115250284	47 kDa	0	0	1	1	1
301	cell surface protein	630_gi|115251820	78 kDa	0	0	0	1	1
303	cell surface protein	630_gi|115251246	51 kDa	0	0	0	1	1
321	putative phage-related cell wall hydrolase (endolysin)	630_gi|115250945	29 kDa	0	0	1	0	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	38 protiens m=17, r=3, h=7, R=6, s/v=2, t=3			1	2	1	2	1
	SE beads only							
118	ribosome-binding factor A	630_gi|115250346	14 kDa	0	0	3	4	0
137	conserved hypothetical protein	630_gi|115249020	12 kDa	0	0	3	3	0

138	BASYS03269, 3479118-3476842 (CounterClockwise) Hypothetical Protein BASYS03269 putative sugar-phosphate isomerase	sta_BASYS03269 115252101	80 kDa	0	0	0	8	0
211	RNA polymerase sigma factor (sigma-43)	630_gi|115250496	44 kDa	0	0	1	3	0
218	50S ribosomal protein L23	630_gi|115249079	11 kDa	0	0	4	2	0
237	BASYS03073, 3241352-3239751 (CounterClockwise) Hypothetical Protein BASYS03073 surface surface protein	stb_BASYS03073 115251847	58 kDa	0	0	2	2	0
242	conserved hypothetical protein	630_gi|115250343	11 kDa	0	0	1	1	0
263	putative thiol peroxidase (bacterioferritin comigratory protein)	630_gi|115250867	18 kDa	0	0	1	3	0
300	putative histidinol-phosphate aminotransferase	630_gi|115250590	40 kDa	0	0	1	1	0
304	putative FAD-binding subunit of oxidoreductase	630_gi|115251155	30 kDa	0	0	1	1	0
306	putative mannosyl-glycoprotein endo-beta-N-acetylglucosamidase	630_gi|115250339	66 kDa	0	0	3	0	0
320	GntR-family transcriptional regulator	630_gi|115250658	14 kDa	0	0	1	1	0
327	putative 16S rRNA processing protein	630_gi|115250289	20 kDa	0	0	1	1	0
328	putative tellurium resistance protein	630_gi|115250693	22 kDa	0	0	1	1	0
334	conserved hypothetical protein	630_gi|115252624	16 kDa	0	0	0	2	0
335	conserved hypothetical protein	630_gi|115252730	7 kDa	0	0	0	2	0
352	RNA polymerase sigma-B factor	630_gi|115249014	30 kDa	0	0	1	1	0
353	penicillinase repressor	630_gi|115249484	15 kDa	0	0	1	1	0
367	putative ribosomal protein	630_gi|115250344	11 kDa	0	0	0	2	0
376	putative signaling protein	630_gi|115250887	132 kDa	0	0	1	1	0
377	putative antibiotic resistance ABC transporter,ATP-binding protein	630_gi|115251650	63 kDa	0	0	0	1	0
403	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898 cell surface protein (S-layer precursor protein)	stt_BASYS02898 115251846	80 kDa	0	0	0	2	0
410	conserved hypothetical protein	630_gi|115249575	22 kDa	0	0	0	1	0
412	ABC transporter, permease protein	630_gi|115250021	30 kDa	0	0	1	0	0
418	bifunctional protein includes: UDP-N-acetylglucosamine pyrophosphorylase and glucosamine-1-phosphate N-acetyltransferase	630_gi|115252576	50 kDa	0	0	1	0	0
419	orotidine 5'-phosphate decarboxylase	630_gi|115252656	26 kDa	0	0	1	0	0
420	queuine tRNA-ribosyltransferase	630_gi|115251855	43 kDa	0	0	0	1	0
442	putative tRNA binding protein	630_gi|115251499	17 kDa	0	0	1	0	0
445	ABC transporter, ATP-binding protein	630_gi|115251001	25 kDa	0	0	1	0	0

471	ATP-dependent Clp protease	630_gi|115249029	91 kDa	0	0	1	0	0
472	conserved hypothetical protein	630_gi|115249281	54 kDa	0	0	1	0	0
473	V-type sodium ATP synthase subunit D	630_gi|115252010	26 kDa	0	0	1	0	0
474	aspartate-semialdehyde dehydrogenase	630_gi|115252281	37 kDa	0	0	1	0	0
477	putative capsular polysaccharide biosynthesis glycosyl transferase	630_gi|115251823	41 kDa	0	0	0	1	0
478	putative acetyltransferase	630_gi|115252032	17 kDa	0	0	0	1	0
521	anti-sigma-B factor antagonist	630_gi|115249012	12 kDa	0	0	0	1	0
524	potassium-transporting ATPase C chain	630_gi|115250634	22 kDa	0	0	1	0	0
156	BASYS00015, 10793-10440 (CounterClockwise) Hypothetical Protein BASYS00015 -	Stbphage1_BASYS00015 -	14 kDa	(1)	0	3	2	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	113 protiens m=68, r=0, h=20, R=12, s/v=9, t=4			1	2	1	2	1
	Crude only							
244	trehalose-6-phosphate hydrolase	630_gi|115252145	66 kDa	0	0	0	0	5
245	alanine racemase	630_gi|115252523	43 kDa	0	0	0	0	5
265	polyribonucleotide nucleotidyltransferase	630_gi|115250354	78 kDa	0	0	0	0	3
280	putative nitroreductase	630_gi|115251628	25 kDa	0	0	0	0	4
281	BASYS00004, 3075-2170 (CounterClockwise) ATPase Of -	stbphage2_BASYS00004 -	33 kDa	0	0	0	0	4
282	putative carbon-nitrogen hydrolase	630_gi|115251789	31 kDa	0	0	0	0	3
283	HPr(Ser) kinase/phosphorylase	630_gi|115252469	36 kDa	0	0	0	0	3
284	methenyltetrahydrofolate cyclohydrolase	630_gi|115249736	23 kDa	0	0	0	0	3
285	BASYS03019, 3175744-3173333 (CounterClockwise) Hypothetical Protein BASYS03019 cell surface protein (putative cell surface-associated cysteine protease)	stb_BASYS03019 115251840	87 kDa	0	0	0	0	2
286	conserved hypothetical protein	630_gi|115252732	22 kDa	0	0	0	0	2
307	putative RNA-binding protein	630_gi|115250347	36 kDa	0	0	0	0	1
308	elongation factor P	630_gi|115250279	21 kDa	0	0	0	0	3
309	tpiA, 3926737-3925994 (CounterClockwise) Hypothetical Protein tpiA triosephosphate isomerase	sta_BASYS03660 115252229	27 kDa	0	0	0	0	3
310	arginyl-tRNA synthetase	630_gi|115249727	65 kDa	0	0	0	0	3
311	uracil-DNA glycosylase	630_gi|115251535	26 kDa	0	0	0	0	3
323	ABC transporter, ATP-binding protein	630_gi|115252673	27 kDa	0	0	0	0	2
329	lysyl-tRNA synthetase	630_gi|115252615	58 kDa	0	0	0	0	3
330	cinserved hypothetical protein	630_gi|115251677	47 kDa	0	0	0	0	3
331	cell surface protein	630_gi|115251849	67 kDa	0	0	0	0	2

332	conserved hypothetical protein	630_gi|115249635	9 kDa	0	0	0	0	2
333	V-type sodium ATP synthase subunit B	630_gi|115252011	51 kDa	0	0	0	0	2
338	tellurium resistance protein	630_gi|115250675	21 kDa	0	0	0	0	1
339	conserved hypothetical protein	630_gi|115249879	33 kDa	0	0	0	0	1
363	chaperone	630_gi|115251074	98 kDa	0	0	0	0	3
364	adenylate kinase	630_gi|115249098	24 kDa	0	0	0	0	2
365	putative lipoprotein	630_gi|115249561	23 kDa	0	0	0	0	2
366	DNA polymerase I	630_gi|115250159	101 kDa	0	0	0	0	2
368	putative glutaminyl-tRNA synthetase	630_gi|115251113	64 kDa	0	0	0	0	2
369	conserved hypothetical protein	630_gi|115251678	10 kDa	0	0	0	0	2
370	dihydrodipicolinate synthase	630_gi|115252280	32 kDa	0	0	0	0	2
371	putative endopeptidase; pz-peptidase	630_gi|115252325	69 kDa	0	0	0	0	2
372	hypothetical protein	630_gi|115251768	78 kDa	0	0	0	0	2
373	ABC transporter, ATP-binding protein	630_gi|115251122	62 kDa	0	0	0	0	2
374	putative peptidyl-prolyl isomerase	630_gi|115250393	28 kDa	0	0	0	0	2
375	putative cyclase	630_gi|115251016	24 kDa	0	0	0	0	2
378	putative ATP/GTP-binding protein	630_gi|115249123	25 kDa	0	0	0	0	1
380	two-component response regulator	630_gi|115250828	28 kDa	0	0	0	0	1
381	inosine-5'-monophosphate dehydrogenase	630_gi|115251390	55 kDa	0	0	0	0	1
382	histidine triad nucleotide-binding protein	630_gi|115251501	13 kDa	0	0	0	0	1
383	ferredoxin	630_gi|115252670	6 kDa	0	0	0	0	1
401	flagellar hook protein	630_gi|115249263	35 kDa	0	0	0	0	2
402	pyrroline-5-carboxylate reductase	630_gi|115250536	29 kDa	0	0	0	0	2
404	mrsA, 259953-258547 (CounterClockwise) Hypothetical Protein mrsA phosphoglucomutase/phosphomannomutase mutase	stb_BASYS00277 -	51 kDa	0	0	0	0	2
405	probable amino-acid ABC transporter,substrate-binding protein	630_gi|115251230	30 kDa	0	0	0	0	2
406	transcription elongation factor	630_gi|115252616	18 kDa	0	0	0	0	2
407	cell surface protein	630_gi|115249861	34 kDa	0	0	0	0	2
413	putative aconitase/3-isopropylmalate dehydratase	630_gi|115249850	69 kDa	0	0	0	0	1
422	bifunctional purine biosynthesis protein includes: phosphoribosylaminoimidazolecarboxamide formyltransferase and IMP cyclohydrolase	630_gi|115249231	57 kDa	0	0	0	0	1
423	putative membrane protein	630_gi|115249537	42 kDa	0	0	0	0	1
424	trans-2-enoyl-ACP reductase	630_gi|115250213	33 kDa	0	0	0	0	1
425	putative ATP phosphoribosyltransferase regulatory subunit	630_gi|115250588	37 kDa	0	0	0	0	1
426	putative imidazole glycerol phosphate synthase subunit	630_gi|115250592	23 kDa	0	0	0	0	1
427	putative phosphoglucomutase	630_gi|115250617	64 kDa	0	0	0	0	1

428	putative membrane protein	630_gi|115250814	41 kDa	0	0	0	0	1
429	putative nitrite and sulfite reductase subunit	630_gi|115250842	34 kDa	0	0	0	0	1
430	SOS regulatory protein	630_gi|115250977	24 kDa	0	0	0	0	1
431	UDP-N-acetylmuramoylalanine--D-glutamate ligase	630_gi|115251704	50 kDa	0	0	0	0	1
432	PTS system, glucose-specific IIa component	630_gi|115252082	17 kDa	0	0	0	0	1
433	PTS system, IIabc component	630_gi|115252143	52 kDa	0	0	0	0	1
434	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	630_gi|115252228	56 kDa	0	0	0	0	1
435	putative nitroreductase	630_gi|115252261	22 kDa	0	0	0	0	1
436	ATP synthase subunit gamma	630_gi|115252529	32 kDa	0	0	0	0	1
437	rffG, 606774-607757 (Clockwise) Hypothetical Protein rffG dTDP-glucose 4,6-dehydratase	sta_BASYS00584 -	38 kDa	0	0	0	0	1
438	conserved hypothetical protein	630_gi|115249636	8 kDa	0	0	0	0	1
439	BASYS01927, 2014266-2015093 (Clockwise) Hypothetical -	sta_BASYS01927 -	33 kDa	0	0	0	0	1
441	cell surface protein	630_gi|115251837	73 kDa	0	0	0	0	1
443	transcription antitermination protein	630_gi|115249064	20 kDa	0	0	0	0	1
447	transketolase	630_gi|115251377	29 kDa	0	0	0	0	1
479	putative aldo/keto reductase	630_gi|115249577	38 kDa	0	0	0	0	1
480	toxin B	630_gi|115249675	270 kDa	0	0	0	0	1
481	NH3-dependent NAD(+) synthetase	630_gi|115249811	28 kDa	0	0	0	0	1
482	putative oxidative stress protein	630_gi|115249845	53 kDa	0	0	0	0	1
483	putative homocitrate/2-isopropylmalate synthase	630_gi|115249849	44 kDa	0	0	0	0	1
484	putative formyltransferase	630_gi|115249858	44 kDa	0	0	0	0	1
485	GntR-family transcriptional regulator	630_gi|115249901	25 kDa	0	0	0	0	1
486	glucosamine-6-phosphate deaminase	630_gi|115250032	28 kDa	0	0	0	0	1
487	fatty acid/phospholipid synthesis protein	630_gi|115250211	37 kDa	0	0	0	0	1
488	putative GTP cyclohydrolase I	630_gi|115250490	21 kDa	0	0	0	0	1
489	conserved hypothetical protein	630_gi|115250679	60 kDa	0	0	0	0	1
490	putative NADPH-dependent FMN reductase	630_gi|115250716	23 kDa	0	0	0	0	1
491	inosine-uridine preferring nucleoside hydrolase	630_gi|115250725	36 kDa	0	0	0	0	1
492	tellurium resistance protein	630_gi|115250845	21 kDa	0	0	0	0	1
493	cytidylate kinase	630_gi|115250861	24 kDa	0	0	0	0	1
494	RpiR-family transcriptional regulator	630_gi|115251102	33 kDa	0	0	0	0	1
495	fructose-6-phosphate aldolase 2	630_gi|115251164	25 kDa	0	0	0	0	1
496	N-acetylneuraminate lyase	630_gi|115251294	33 kDa	0	0	0	0	1
497	conserved hypothetical protein	630_gi|115251601	68 kDa	0	0	0	0	1
498	guanylate kinase	630_gi|115251645	23 kDa	0	0	0	0	1
499	probable peptidase	630_gi|115251712	48 kDa	0	0	0	0	1

500	PTS system, glucose-specific IIa component	630_gi|115251717	18 kDa	0	0	0	0	1
501	hypothetical protein	630_gi|115251805	26 kDa	0	0	0	0	1
502	PTS system, IIa component	630_gi|115252072	17 kDa	0	0	0	0	1
503	putative D-isomer specific 2-hydroxyacid dehydrogenase	630_gi|115252366	36 kDa	0	0	0	0	1
504	porphobilinogen deaminase	630_gi|115252481	33 kDa	0	0	0	0	1
506	yieK, 1939575-1940309 (Clockwise) Hypothetical Protein yieK -	stb_BASYS01888 -	28 kDa	0	0	0	0	1
507	rffG, 3225623-3224601 (CounterClockwise) Hypothetical Protein rffG -	stb_BASYS03062 -	39 kDa	0	0	0	0	1
511	aspartate carbamoyltransferase catalytic chain	630_gi|115249194	35 kDa	0	0	0	0	1
513	radical SAM-superfamily protein	630_gi|115250190	71 kDa	0	0	0	0	1
522	conserved hypothetical protein	630_gi|115251487	22 kDa	0	0	0	0	1
527	anaerobic ribonucleoside-triphosphate reductase	630_gi|115249116	89 kDa	0	0	0	0	1
529	conserved hypothetical protein	630_gi|115251463	32 kDa	0	0	0	0	1
530	BASYS02387, 2486053-2485664 (CounterClockwise) Hypothetical Protein AGRC conserved hypothetical protein	sta_BASYS02387 115250980	15 kDa	0	0	0	0	1
532	adenylosuccinate synthetase	630_gi|115252719	47 kDa	0	0	0	0	1
535	putative sigma 54 modulation protein	630_gi|115251497	22 kDa	0	0	0	0	1
542	PTS system, IId component	630_gi|115249298	31 kDa	0	0	0	0	1
543	putative L-asparaginase	630_gi|115251569	36 kDa	0	0	0	0	1
545	BASYS01697, 1747000-1746578 (CounterClockwise) Hypothetical Protein BASYS01697 -	stb_BASYS01697 -	16 kDa	0	0	0	0	1
549	ABC transporter, ATP-binding protein	630_gi|115249802	59 kDa	0	0	0	0	1
550	LacI-family transcriptional regulator	630_gi|115249813	37 kDa	0	0	0	0	1
553	cell surface protein	630_gi|115249450	41 kDa	0	0	0	0	1
555	putative flavodoxin	630_gi|115251878	20 kDa	0	0	0	0	1
556	conserved hypothetical protein	630_gi|115250439	23 kDa	0	0	0	0	1
557	putative FolD bifunctional protein includes: methylenetetrahydrofolate dehydrogenase; methenyltetrahydrofolate cyclohydrolase	630_gi|115249737	31 kDa	0	0	0	0	1

Table A7 – *Proteins identified in strain Tra5/5 ProteoMiner treated broth extracts by 1D GE followed by LC-MS/MS analysis with identification number, predicted mass and number of unique peptides detected in each sample. The number of proteins of different types are show as follows; metabolic (m), surface and virulence (s/v), hypothetical (h), regulatory (R), transporter (t) and ribosomal (r).*

	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	200 proteins m=118, r=41, h=17, R=15, s/v=6, t=3			1	2	1	2	1
	All samples							
	acyl-CoA dehydrogenase, short-chain specific	630_gi|115249405	41 kDa	23	21	21	16	15
2	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	630_gi|115249401	44 kDa	22	23	15	10	20
3	acetyl-CoA acetyltransferase	630_gi|115250080	41 kDa	14	13	15	15	26
4	30S ribosomal protein S7	630_gi|115249073	18 kDa	12	11	10	10	11
5	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249403	47 kDa	19	19	12	10	16
6	electron transfer flavoprotein alpha-subunit	630_gi|115249407	37 kDa	17	15	13	12	11
7	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249404	42 kDa	24	23	17	17	19
8	electron transfer flavoprotein beta-subunit	630_gi|115249406	29 kDa	14	14	11	10	10
9	putative rubrerythrin	630_gi|115250565	20 kDa	15	15	10	9	10
10	pyruvate-flavodoxin oxidoreductase	630_gi|115251733	129 kDa	33	31	27	21	22
11	NAD-specific glutamate dehydrogenase	630_gi|115249189	46 kDa	18	17	11	10	21
12	putative amino acid aminotransferase	630_gi|115252729	45 kDa	21	19	14	12	19
13	30S ribosomal protein S3	630_gi|115249083	30 kDa	11	11	13	11	2
14	30S ribosomal protein S19	630_gi|115249081	11 kDa	5	5	5	4	3
15	elongation factor TU	630_gi|115249061	44 kDa	18	16	11	9	11
16	30S ribosomal protein S4	630_gi|115249105	24 kDa	11	10	8	8	4
17	50S ribosomal protein L1	630_gi|115249066	25 kDa	10	10	9	9	6
18	30S ribosomal protein S5	630_gi|115249094	18 kDa	6	6	8	9	3
19	50S ribosomal protein L24	630_gi|115249088	11 kDa	4	4	5	5	1
20	50S ribosomal protein L19	630_gi|115250291	13 kDa	3	4	8	8	2
21	50S ribosomal protein L13	630_gi|115249112	16 kDa	8	7	7	7	2
22	30S ribosomal protein S9	630_gi|115249113	15 kDa	7	10	9	8	5
23	adhE, 3383815-3381161 (CounterClockwise) Hypothetical Protein adhE aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase	stb_BASYS03189 115252023	97 kDa	17	17	16	16	26

24	rpsJ, 420934-421245 (Clockwise) Hypothetical Protein rpsJ 30S ribosomal protein S10	sta_BASYS00411 115249076	12 kDa	7	7	5	5	4
25	butyryl-CoA dehydrogenase	630_gi|115250075	41 kDa	13	15	15	13	15
26	50S ribosomal protein L4	630_gi|115249078	24 kDa	10	8	9	7	5
27	putative ruberythrin	630_gi|115250515	20 kDa	6	7	4	4	4
28	30S ribosomal protein S16	630_gi|115250287	10 kDa	3	3	4	4	3
29	50S ribosomal protein L21	630_gi|115250193	11 kDa	4	3	3	3	2
30	50S ribosomal protein L2	630_gi|115249080	30 kDa	7	7	8	8	3
31	30S ribosomal protein S8	630_gi|115249091	15 kDa	5	5	6	6	6
32	30S ribosomal protein S20	630_gi|115251527	10 kDa	2	2	2	2	2
33	50S ribosomal protein L5	630_gi|115249089	20 kDa	10	9	11	11	6
34	putative oxidoreductase, thiamine diP-binding subunit	630_gi|115249125	39 kDa	10	13	8	7	3
35	eutG, 2931498-2930332 (CounterClockwise) Hypothetical Protein eutG NAD-dependent 4-hydroxybutyrate dehydrogenase	sta_BASYS02789 115251393	44 kDa	11	11	7	3	6
36	glycine/sarcosine/betaine reductase complex component C beta subunit	630_gi|115251404	55 kDa	10	11	9	9	13
37	50S ribosomal protein L6	630_gi|115249092	20 kDa	5	4	8	9	1
38	50S ribosomal protein L15	630_gi|115249096	16 kDa	5	7	7	7	3
39	aspartate aminotransferase	630_gi|115250375	45 kDa	15	14	9	8	17
40	succinate-semialdehyde dehydrogenase NAD(P)+	630_gi|115251397	51 kDa	15	14	15	12	17
41	electron transfer flavoprotein alpha-subunit	630_gi|115250077	36 kDa	12	11	8	8	13
42	electron transfer flavoprotein beta-subunit	630_gi|115250076	29 kDa	8	10	4	2	14
43	3-hydroxybutyryl-CoA dehydrogenase	630_gi|115250079	31 kDa	7	7	6	5	12
44	(R)-2-hydroxyisocaproate dehydrogenase	630_gi|115249400	37 kDa	10	10	8	5	12
45	50S ribosomal protein L3	630_gi|115249077	22 kDa	12	12	11	10	5
46	30S ribosomal protein S11	630_gi|115249104	14 kDa	4	3	4	3	2
47	50S ribosomal protein L17	630_gi|115249107	13 kDa	2	2	3	2	1
48	cell surface protein	630_gi|115251764	72 kDa	7	10	16	17	2
49	glycine reductase complex component B alpha and beta subunits	630_gi|115251407	46 kDa	8	8	6	4	14
50	stage 0 sporulation protein A	630_gi|115250247	31 kDa	6	4	12	10	4
51	gamma-aminobutyrate metabolism dehydratase/isomerase includes: 4-hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta-isomerase	630_gi|115251396	55 kDa	8	9	7	6	14
52	30S ribosomal protein S13	630_gi|115249103	14 kDa	6	7	5	4	3
55	fliC, 615650-616609 (Clockwise) Hypothetical Protein fliC flagellin subunit	sta_BASYS00598 115249247	34 kDa	8	7	8	9	2
56	rubrerythrin	630_gi|115249842	21 kDa	4	4	2	2	6

57	DNA-directed RNA polymerase beta chain	630_gi|115249070	139 kDa	19	21	21	10	20
58	ferredoxin	630_gi|115249124	8 kDa	5	5	5	5	6
59	pepD, 849602-848151 (CounterClockwise) Hypothetical Protein pepD putative aminoacyl-histidine dipeptidase	stt_BASYS00847 115249724	53 kDa	10	10	4	3	10
60	50S ribosomal protein L14	630_gi|115249087	13 kDa	5	4	4	4	1
61	translation initiation factor IF-3	630_gi|115249701	19 kDa	5	5	11	13	1
62	50S ribosomal protein L29	630_gi|115249085	8 kDa	3	3	4	3	2
63	50S ribosomal protein L27	630_gi|115250195	10 kDa	3	4	5	6	1
64	50S ribosomal protein L16	630_gi|115249084	16 kDa	2	2	3	2	1
65	putative proline racemase	630_gi|115252294	36 kDa	10	10	7	5	11
66	50S ribosomal protein L22	630_gi|115249082	12 kDa	5	5	6	6	4
67	formate--tetrahydrofolate ligase	630_gi|115249735	60 kDa	3	2	5	6	20
68	putative oxidoreductase, acetyl-CoA synthase subunit	630_gi|115249184	68 kDa	11	13	10	8	5
69	DNA-directed RNA polymerase beta' chain	630_gi|115249071	130 kDa	7	11	21	11	9
70	50S ribosomal protein L10	630_gi|115249067	19 kDa	7	7	6	6	6
71	norV, 2106146-2108677 (Clockwise) Hypothetical Protein norV putative nitric oxide reductase flavoprotein	sta_BASYS02024 115250664	94 kDa	8	11	3	1	20
72	translation elongation factor G	630_gi|115249074	76 kDa	11	13	7	5	19
73	putative formate acetyltransferase	630_gi|115252338	89 kDa	15	15	6	3	12
74	putative propanediol utilization protein	630_gi|115251734	20 kDa	3	3	4	3	6
75	probable peptidase	630_gi|115251664	40 kDa	9	9	4	1	3
76	enolase	630_gi|115252227	46 kDa	8	8	2	2	9
77	indolepyruvate oxidoreductase subunit	630_gi|115251433	66 kDa	9	8	5	4	19
78	glyceraldehyde-3-phosphate dehydrogenase 2	630_gi|115252231	36 kDa	6	6	7	6	6
79	glyS, 2611727-2609661 (CounterClockwise) Hypothetical Protein glyS	stt_BASYS02514 115251485	78 kDa	9	16	16	11	9
422	glycyl-tRNA synthetase beta chain glyS, 2743525-2741459 (CounterClockwise) Hypothetical Protein glyS glycyl-tRNA synthetase beta chain	stb_BASYS02649 115251485	78 kDa	0	0	1	1	0
80	putative dinitrogenase iron-molybdenum cofactor	630_gi|115250736	13 kDa	3	4	2	2	4
81	glycine/sarcosine/betaine reductase complex component C alpha subunit	630_gi|115251403	41 kDa	8	7	6	6	10
82	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	630_gi|115249053	55 kDa	8	8	4	3	2
83	isoleucyl-tRNA synthetase	630_gi|115251669	120 kDa	11	10	7	2	6
84	50S ribosomal protein L7/L12	630_gi|115249068	13 kDa	2	2	2	2	2

85	BASYS02898, 3050518-3048248 (CounterClockwise) Hypothetical Protein BASYS02898 cell surface protein (S-layer precursor protein)	stt_BASYS02898 115251846	80 kDa	3	2	2	4	27
86	30S ribosomal protein S15	630_gi|115250352	10 kDa	5	5	5	2	3
87	50S ribosomal protein L36	630_gi|115249102	4 kDa	1	1	1	1	1
88	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex,methyltransferase subunit	630_gi|115249744	29 kDa	8	8	3	1	8
89	putative anaerobic nitric oxide reductase flavorubredoxin	630_gi|115250189	45 kDa	8	7	2	1	1
90	putative subunit of oxidoreductase	630_gi|115249127	20 kDa	4	3	2	2	2
91	30S ribosomal protein S2	630_gi|115251194	27 kDa	6	7	7	6	1
92	chaperone protein	630_gi|115251515	66 kDa	1	1	1	2	15
93	60 kDa chaperonin	630_gi|115249204	58 kDa	9	8	2	5	11
94	conserved hypothetical protein	630_gi|115249288	15 kDa	1	2	2	3	4
95	putative phosphate butyryltransferase	630_gi|115249732	33 kDa	4	5	7	5	8
96	3-hydroxybutyryl-CoA dehydratase	630_gi|115250078	28 kDa	3	3	1	1	8
98	PTS system, IIb component	630_gi|115252071	11 kDa	2	2	3	4	3
99	putative NUDIX-family hydrolase	630_gi|115249807	17 kDa	2	3	6	5	7
100	threonine dehydratase catabolic	630_gi|115251567	43 kDa	9	8	8	6	4
101	putative RNA-binding protein	630_gi|115249151	21 kDa	6	5	5	4	2
102	50S ribosomal protein L9	630_gi|115252722	17 kDa	3	3	9	9	2
103	activator of 2-hydroxyisocaproyl-CoA dehydratase	630_gi|115249402	29 kDa	5	5	3	3	5
104	glycine reductase complex component B gamma subunit	630_gi|115251405	47 kDa	3	2	2	2	12
105	threonyl-tRNA synthetase	630_gi|115249589	74 kDa	5	8	5	3	1
107	50S ribosomal protein L11	630_gi|115249065	15 kDa	1	2	1	2	1
108	putative translation inhibitor endoribonuclease	630_gi|115251566	14 kDa	1	1	1	1	1
109	conserved hypothetical protein	630_gi|115251678	10 kDa	2	2	1	1	6
110	putative glutamate synthase NADPH small chain	630_gi|115250578	50 kDa	7	6	1	4	14
112	trigger factor	630_gi|115252362	48 kDa	6	5	5	5	5
113	conserved hypothetical protein	630_gi|115252289	45 kDa	4	4	3	4	4
114	proline reductase subunit proprotein	630_gi|115252300	68 kDa	5	5	3	1	10
115	formate acetyltransferase	630_gi|115249776	84 kDa	7	9	2	0	15
116	oligopeptide ABC transporter, substrate-binding lipoprotein	630_gi|115249872	59 kDa	6	6	4	3	11
117	50S ribosomal protein L20	630_gi|115249703	13 kDa	3	3	2	1	2
118	4-hydroxybutyrate CoA transferase	630_gi|115251394	48 kDa	3	3	3	1	12
119	tellurium resistance protein	630_gi|115250846	21 kDa	3	3	7	7	4
121	putative 2-hydroxyacyl-CoA dehydratase	630_gi|115250793	44 kDa	3	3	2	2	1

122	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	630_gi|115249740	49 kDa	3	2	4	3	15
123	putative universal stress protein	630_gi|115249829	15 kDa	4	6	3	2	7
124	DNA-directed RNA polymerase alpha chain	630_gi|115249106	35 kDa	6	9	6	7	2
125	leucyl-tRNA synthetase	630_gi|115251575	92 kDa	6	8	3	0	6
128	putative ribosomal protein	630_gi|115249100	11 kDa	4	4	5	5	2
129	putative translation elongation factor	630_gi|115249025	72 kDa	4	7	6	3	8
130	peptidase	630_gi|115251315	68 kDa	6	9	4	2	3
132	conserved hypothetical protein	630_gi|115252075	17 kDa	3	2	2	2	2
133	putative indolepyruvate oxidoreductase subunit	630_gi|115251432	21 kDa	3	3	0	1	4
134	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, alpha subunit	630_gi|115249743	50 kDa	2	2	5	3	7
135	ribose-phosphate pyrophosphokinase	630_gi|115252575	34 kDa	6	6	3	2	1
136	abgA, 2922374-2921253 (CounterClockwise) Hypothetical Protein abgA putative amidohydrolase	stt_BASYS02795 115251748	41 kDa	5	7	1	0	5
137	GTP-sensing transcriptional pleiotropic repressor	630_gi|115250309	29 kDa	4	6	5	5	5
138	DNA-binding protein HU	630_gi|115252557	10 kDa	2	2	1	1	3
140	ATP synthase beta chain	630_gi|115252528	50 kDa	4	5	3	2	4
141	metG, 3942695-3940758 (CounterClockwise) Hypothetical Protein metG methionyl-tRNA synthetase	stt_BASYS03683 115252601	?	6	8	2	0	3
142	AsnC-family transcriptional regulator	630_gi|115252605	16 kDa	2	3	1	1	6
143	proline reductase	630_gi|115252298	26 kDa	1	1	1	1	4
144	ATP synthase alpha chain	630_gi|115252530	55 kDa	4	3	2	2	3
145	tellurium resistance protein	630_gi|115250676	22 kDa	1	1	2	1	4
148	30S ribosomal protein S6	630_gi|115252728	11 kDa	2	3	1	1	3
150	conserved hypothetical protein	630_gi|115251421	12 kDa	2	2	1	1	3
152	asparaginyl-tRNA synthetase	630_gi|115251299	54 kDa	5	8	1	0	2
154	V-type sodium ATP synthase subunit A	630_gi|115252012	66 kDa	1	1	1	0	10
157	thioredoxin reductase	630_gi|115251409	34 kDa	3	4	0	1	4
160	proC, 4066107-4065304 (CounterClockwise) Hypothetical Protein proC pyrroline-5-carboxylate reductase	sta_BASYS03787 115252337	28 kDa	6	7	1	1	1
161	conserved hypothetical protein	630_gi|115249575	22 kDa	1	1	2	1	1
162	putative decarboxylase	630_gi|115251040	12 kDa	2	2	1	0	7
163	putative serine hydroxymethyltransferase	630_gi|115251778	46 kDa	4	3	4	2	2
164	transketolase	630_gi|115251376	32 kDa	4	5	1	4	4

167	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein	630_gi|115250698	92 kDa	3	3	4	1	7
168	putative fructose-bisphosphate aldolase	630_gi|115249409	33 kDa	5	5	2	1	1
172	valyl-tRNA synthetase	630_gi|115252312	103 kDa	4	6	2	1	1
175	ABC transporter, substrate-binding lipoprotein	630_gi|115249887	36 kDa	4	3	1	2	6
176	butyrate kinase	630_gi|115249122	39 kDa	2	2	2	4	6
177	stage V sporulation protein G	630_gi|115252577	10 kDa	2	2	1	0	2
179	cysteinyl-tRNA synthetase	630_gi|115249055	54 kDa	1	2	2	4	6
181	putative DNA-binding protein	630_gi|115249180	23 kDa	2	3	5	4	3
183	glycogen synthase	630_gi|115249898	56 kDa	3	4	1	1	1
185	thioredoxin	630_gi|115250733	12 kDa	2	2	1	1	3
187	putative dehydrogenase, electron transfer subunit	630_gi|115250577	33 kDa	4	4	1	0	5
191	conserved hypothetical protein	630_gi|115250551	24 kDa	2	3	3	3	2
192	putative 5-nitroimidazole reductase	630_gi|115250500	18 kDa	3	4	1	1	2
194	hypothetical protein	630_gi|115249824	14 kDa	1	2	1	1	1
197	putative transcriptional regulator	630_gi|115250535	23 kDa	2	2	2	2	2
198	putative 16S rRNA processing protein	630_gi|115250289	20 kDa	4	4	3	2	2
199	succinyl-CoA:coenzyme A transferase	630_gi|115251398	56 kDa	2	3	2	0	2
200	conserved hypothetical protein	630_gi|115250197	14 kDa	3	4	2	0	2
201	oligopeptide ABC transporter, substrate-binding protein	630_gi|115251723	58 kDa	1	1	1	1	8
204	pyruvate carboxylase	630_gi|115249024	129 kDa	2	2	7	2	1
205	conserved hypothetical protein	630_gi|115251576	13 kDa	1	1	2	2	2
206	ribosome recycling factor	630_gi|115251191	21 kDa	0	1	2	3	4
207	DNA gyrase subunit A	630_gi|115249009	91 kDa	1	1	5	5	1
210	adenylosuccinate lyase	630_gi|115250370	55 kDa	2	2	3	2	1
211	elongation factor Ts	630_gi|115251193	33 kDa	3	4	1	1	3
212	conserved hypothetical protein	630_gi|115251617	13 kDa	3	2	1	1	1
218	ATP synthase subunit delta	630_gi|115252531	21 kDa	2	1	3	3	2
219	conserved hypothetical protein	630_gi|115250989	11 kDa	2	2	0	1	2
221	BASYS02853, 2966343-2964979 (CounterClockwise) Hypothetical Protein BASYS02853 cinserved hypothetical protein	stb_BASYS02853 115251677	48 kDa	2	1	1	0	8
226	PTS system, IIb component	630_gi|115249296	18 kDa	0	1	1	1	3
229	SpoIIIJ-associated protein	630_gi|115252741	24 kDa	2	4	1	1	1
230	seryl-tRNA synthetase	630_gi|115249017	49 kDa	1	1	3	1	4
231	conserved hypothetical protein	630_gi|115250233	13 kDa	1	1	1	1	2
232	putative acetyltransferase	630_gi|115249853	19 kDa	1	2	1	0	3
241	uracil phosphoribosyltransferase	630_gi|115252539	23 kDa	3	3	3	3	2
242	phosphoribosylaminoimidazole carboxylase catalytic subunit	630_gi|115249226	17 kDa	2	2	2	1	2
244	lysyl-tRNA synthetase	630_gi|115252615	58 kDa	0	1	3	2	2
245	glycine cleavage system P protein	630_gi|115250699	54 kDa	1	1	1	1	1
248	pyruvate kinase	630_gi|115252454	63 kDa	1	1	1	0	4
250	aspartate aminotransferase	630_gi|115249115	44 kDa	2	2	2	1	1
251	putative regulatory protein	630_gi|115251119	15 kDa	1	1	2	1	1

274	putative 30S ribosomal protein S1	630_gi|115250007	48 kDa	2	2	1	3	1
275	hypothetical protein	630_gi|115250298	10 kDa	1	1	1	0	2
276	transketolase	630_gi|115251377	29 kDa	1	1	1	1	1
289	GntR-family transcriptional regulator	630_gi|115250423	24 kDa	1	1	3	2	1
290	D-alanine--poly(phosphoribitol) ligase subunit 2 (D-alanyl carrier protein)	630_gi|115251904	9 kDa	2	2	1	0	2
291	putative aspartyl-tRNA synthetase	630_gi|115251791	67 kDa	1	2	0	2	3
293	tellurium resistance protein	630_gi|115250677	23 kDa	2	2	1	0	3
294	tyrosyl-tRNA synthetase	630_gi|115250562	46 kDa	2	0	1	1	4
305	putative tellurite resistance protein	630_gi|115250680	44 kDa	2	1	0	3	2
308	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	630_gi|115249733	69 kDa	1	1	1	0	1
318	anti-sigma F factor antagonist	630_gi|115249787	13 kDa	2	2	1	0	1
328	BASYS01206, 1239342-1239710 (Clockwise) Hypothetical Protein BASYS01206 -	stb_BASYS01206 -	14 kDa	0	2	1	0	1
349	putative D-tyrosyl-tRNA protein	630_gi|115251795	16 kDa	1	1	1	0	1
354	conserved hypothetical protein	630_gi|115250343	11 kDa	1	0	1	1	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	67 proteins m=47, r=0, h=12, R=8, s/v=0, t=0			1	2	1	2	1
	Ebeads + crude							
127	acetate kinase	630_gi|115250207	43 kDa	4	2	0	0	11
146	putative cold shock protein	630_gi|115250391	7 kDa	1	1	0	0	1
147	putative aspartate aminotransferase	630_gi|115251434	47 kDa	0	1	0	0	12
149	rubredoxin oxidoreductase (desulfoferrodoxin)	630_gi|115249844	14 kDa	1	2	0	0	4
153	single-strand binding protein	630_gi|115252727	16 kDa	4	3	0	0	5
155	clpP, 3641751-3641167 (CounterClockwise) Hypothetical Protein clpP ATP-dependent Clp protease proteolytic subunit	stt_BASYS03406 115252361	21 kDa	3	3	0	0	8
158	ribulose-phosphate 3-epimerase	630_gi|115251631	24 kDa	2	2	0	0	2
174	putative cyclase	630_gi|115251016	24 kDa	1	0	0	0	8
178	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, small subunit	630_gi|115249742	34 kDa	4	3	0	0	5
180	tpiA, 3685612-3684869 (CounterClockwise) Hypothetical Protein tpiA triosephosphate isomerase	stb_BASYS03462 115252229	27 kDa	2	1	0	0	7
188	NifU-like protein	630_gi|115250314	16 kDa	2	3	0	0	3
195	nitroreductase-family protein	630_gi|115250155	33 kDa	4	3	0	0	4
202	butyrate kinase	630_gi|115251431	39 kDa	2	2	0	0	9
217	putative nitroreductase	630_gi|115251628	25 kDa	1	0	0	0	5

222	ybgG, 3247187-3244587 (CounterClockwise) Hypothetical Protein ybgG putative glycosyl hydrolase	stt_BASYS03063 115252069	100 kDa	3	5	0	0	4
223	ribose-5-phosphate isomerase 2	630_gi|115252540	16 kDa	1	2	0	0	5
224	ferredoxin	630_gi|115252670	6 kDa	1	1	0	0	1
203	BASYS00907, 938664-938939 (Clockwise) Hypothetical Protein BASYS00907 conserved hypothetical protein	sta_BASYS00907 115249602	10 kDa	2	2	1	1	2
227	inosine-uridine preferring nucleoside hydrolase	630_gi|115250725	36 kDa	2	2	0	0	1
233	putative transaldolase	630_gi|115251384	23 kDa	3	2	0	0	3
235	orotate phosphoribosyltransferase	630_gi|115249197	21 kDa	1	1	0	0	3
246	polyribonucleotide nucleotidyltransferase	630_gi|115250354	78 kDa	3	4	0	0	4
247	pheT, 1093175-1095568 (Clockwise) Phenylalanyl-tRNA synthetase beta chain phenylalanyl-tRNA synthetase beta chain	sta_BASYS01042 115249716	89 kDa	3	3	0	0	4
252	conserved hypothetical protein	630_gi|115249812	26 kDa	1	2	0	0	3
253	electron transfer flavoprotein beta-subunit	630_gi|115249821	29 kDa	1	1	0	0	1
259	methenyltetrahydrofolate cyclohydrolase	630_gi|115249736	23 kDa	1	2	0	0	4
266	putative glycine cleavage system H protein	630_gi|115249746	14 kDa	2	1	0	0	2
268	heat shock protein	630_gi|115251516	24 kDa	2	0	0	0	6
270	ribose-5-phosphate isomerase 1	630_gi|115251375	16 kDa	0	1	0	0	5
278	putative peptidase	630_gi|115252582	40 kDa	1	1	0	0	3
281	putative sigma 54 modulation protein	630_gi|115251497	22 kDa	2	4	0	0	1
282	anti-sigma F factor	630_gi|115249788	16 kDa	0	3	0	0	4
285	ferritin	630_gi|115251248	20 kDa	1	1	0	0	1
286	BASYS00512, 475629-475835 (Clockwise) Hypothetical Protein BASYS00512 -	stt_BASYS00512 -	8 kDa	1	1	0	0	1
297	translation initiation factor IF-2	630_gi|115250345	70 kDa	2	3	0	0	1
298	conserved hypothetical protein	630_gi|115252525	17 kDa	1	2	0	0	2
300	nitrilase (carbon-nitrogen hydrolase)	630_gi|115251895	34 kDa	1	1	0	0	2
302	hypothetical protein	630_gi|115249752	14 kDa	1	1	0	0	1
307	putative phosphoribosyltransferase	630_gi|115251742	20 kDa	2	2	0	0	1
319	hypothetical protein	630_gi|115249299	18 kDa	2	1	0	0	1
321	BASYS00476, 460646-460296 (CounterClockwise) Transcriptional Regulator -	stt_BASYS00476 -	14 kDa	1	1	0	0	1
326	conserved hypothetical protein	630_gi|115252038	35 kDa	2	2	0	0	2
332	probable sugar O-acetyltransferase	630_gi|115250709	20 kDa	1	1	0	0	1
334	adenylate kinase	630_gi|115249098	24 kDa	1	0	0	0	3

335	conserved hypothetical protein	630_gi|115250701	12 kDa	0	2	0	0	3
336	putative glycine reductase complex component	630_gi|115251410	14 kDa	0	1	0	0	3
337	hypothetical protein	630_gi|115252299	11 kDa	1	1	0	0	1
350	dihydrodipicolinate synthase	630_gi|115252282	32 kDa	1	1	0	0	2
351	glycine/sarcosine/betaine reductase complex component A	630_gi|115251406	17 kDa	1	1	0	0	1
356	putative iron-sulfur protein	630_gi|115249747	72 kDa	1	0	0	0	3
358	biotin carboxyl carrier protein of acetyl-CoA carboxylase	630_gi|115250985	17 kDa	0	1	0	0	1
363	putative glutaminyl-tRNA synthetase	630_gi|115251113	64 kDa	1	1	0	0	2
372	putative phosphoesterase	630_gi|115250069	21 kDa	1	1	0	0	1
376	conserved hypothetical protein	630_gi|115249350	16 kDa	1	0	0	0	3
377	putative alanine racemase	630_gi|115251674	27 kDa	1	0	0	0	3
378	putative regulatory protein	630_gi|115250392	18 kDa	0	1	0	0	2
379	putative ATP-binding protein	630_gi|115250774	29 kDa	1	1	0	0	0
391	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit	630_gi|115249745	77 kDa	1	1	0	0	1
398	methylglyoxal synthase	630_gi|115250185	15 kDa	0	1	0	0	2
402	phosphopantetheine adenylyltransferase	630_gi|115251613	18 kDa	1	1	0	0	1
415	putative preprotein translocase	630_gi|115252635	19 kDa	1	1	0	0	1
423	putative metal dependent phosphohydrolase	630_gi|115251577	22 kDa	0	2	0	0	1
429	phosphoglycerate kinase	630_gi|115252230	43 kDa	0	1	0	0	2
452	anti-sigma-B factor (serine-protein kinase)	630_gi|115249013	15 kDa	1	1	0	0	1
457	6-phosphofructokinase	630_gi|115252455	34 kDa	1	0	0	0	1
474	hypothetical protein	630_gi|115252039	74 kDa	0	1	0	0	1
500	conserved hypothetical protein	630_gi|115249281	54 kDa	0	1	0	0	1
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		crude
	58 proteins m=34, r=1, h=12, R=9, s/v=1, t=1			1	2	1	2	1
	Ebeads only							
216	ribonuclease R	630_gi|115252220	82 kDa	3	5	0	0	0
261	conserved hypothetical protein	630_gi|115251612	20 kDa	3	4	0	0	0
311	putative 23S RNA and tRNA pseudouridine synthase	630_gi|115251104	27 kDa	1	3	0	0	0
312	glycyl-tRNA synthetase alpha chain	630_gi|115251486	34 kDa	2	3	0	0	0
322	ferric uptake regulation protein	630_gi|115250321	18 kDa	1	3	0	0	0
323	putative RNA-binding mediating protein	630_gi|115249006	8 kDa	1	1	0	0	0
329	conserved hypothetical protein	630_gi|115252618	28 kDa	2	2	0	0	0
352	carbamoyl-phosphate synthase,pyrimidine-specific, large chain	630_gi|115252652	119 kDa	2	3	0	0	0
353	putative tRNA/rRNA methyltransferase	630_gi|115249058	27 kDa	2	2	0	0	0
367	L-serine dehydratase	630_gi|115252279	43 kDa	2	2	0	0	0

368	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252466	68 kDa	2	2	0	0	0
369	conserved hypothetical protein	630_gi|115252583	51 kDa	2	2	0	0	0
388	biotin carboxylase (acetyl-CoA carboxylase subunit A)	630_gi|115250984	51 kDa	2	2	0	0	0
399	putative Xaa-Pro dipeptidase	630_gi|115251402	41 kDa	2	1	0	0	0
400	L-aspartate-beta-decarboxylase	630_gi|115251568	62 kDa	2	1	0	0	0
403	putative regulatory protein	630_gi|115250074	17 kDa	1	1	0	0	0
417	cell division protein	630_gi|115251697	41 kDa	1	2	0	0	0
424	conserved hypothetical protein	630_gi|115250804	33 kDa	1	1	0	0	0
425	conserved hypothetical protein	630_gi|115251493	18 kDa	1	1	0	0	0
426	precorrin-4 C(11)-methyltransferase	630_gi|115252486	28 kDa	1	1	0	0	0
427	carbamoyl-phosphate synthase,pyrimidine-specific, small chain	630_gi|115252653	39 kDa	1	1	0	0	0
442	cfa, 551238-550051 (CounterClockwise) Hypothetical Protein cfa cyclopropane-fatty-acyl-phospholipid synthase	sta_BASYS00534 115249187	46 kDa	1	1	0	0	0
458	putative oxidoreductase, electron transfer subunit	630_gi|115249185	18 kDa	1	1	0	0	0
459	amidophosphoribosyltransferase	630_gi|115249228	51 kDa	1	1	0	0	0
460	radical SAM-superfamily protein	630_gi|115250727	27 kDa	1	1	0	0	0
461	putative molybdopterin biosynthesis protein	630_gi|115251123	37 kDa	1	1	0	0	0
462	putative ribosomal protein L11 methyltransferase	630_gi|115251504	35 kDa	1	1	0	0	0
463	DNA gyrase subunit B	630_gi|115249008	71 kDa	1	1	0	0	0
464	BASYS00297, 294213-293905 (CounterClockwise) Hypothetical Protein BASYS00297 hypothetical protein	sta_BASYS00297 115252720	12 kDa	1	2	0	0	0
476	stage V sporulation protein T	630_gi|115252560	20 kDa	1	1	0	0	0
478	putative acetyltransferase	630_gi|115250244	18 kDa	1	1	0	0	0
480	conserved hypothetical protein	630_gi|115251531	36 kDa	1	1	0	0	0
487	septum site-determining protein (cell division inhibitor)	630_gi|115250182	29 kDa	1	1	0	0	0
488	probable hydrolase	630_gi|115251629	24 kDa	1	1	0	0	0
489	UTP--glucose-1-phosphate uridylyltransferase	630_gi|115251766	33 kDa	1	1	0	0	0
490	glucose inhibited division protein A	630_gi|115252739	71 kDa	1	1	0	0	0
495	chaperone	630_gi|115251074	98 kDa	0	2	0	0	0
496	N-carbamoyl-L-amino acid hydrolase	630_gi|115251081	44 kDa	2	0	0	0	0
499	conserved hypothetical protein	630_gi|115251511	30 kDa	1	1	0	0	0
535	cold shock protein	630_gi|115251036	7 kDa	1	1	0	0	0
537	single-stranded DNA binding protein	630_gi|115252292	25 kDa	0	1	0	0	0
568	peptidase T	630_gi|115250067	45 kDa	1	0	0	0	0

570	putative fructose-1,6-bisphosphatase	630_gi|115250224	76 kDa	0	1	0	0	0
571	conserved hypothetical protein	630_gi|115250280	17 kDa	0	1	0	0	0
572	putative transcriptional regulator	630_gi|115250685	21 kDa	0	1	0	0	0
573	guanine deaminase	630_gi|115250705	48 kDa	0	1	0	0	0
574	selenocysteine-specific elongation factor	630_gi|115251547	72 kDa	0	1	0	0	0
575	cell-division initiation protein	630_gi|115251670	20 kDa	0	1	0	0	0
577	BASYS02498, 2595334-2595041 (CounterClockwise) Hypothetical Protein BASYS02498 hypothetical protein	sta_BASYS02498 115251100	12 kDa	0	2	0	0	0
579	conserved hypothetical protein	630_gi|115250029	32 kDa	0	1	0	0	0
580	conserved hypothetical protein	630_gi|115251206	9 kDa	1	0	0	0	0
589	putative sugar-phosphate kinase	630_gi|115249217	48 kDa	1	0	0	0	0
594	utp--glucose-1-phosphate uridylyltransferase (general stress protein 33)	630_gi|115252549	36 kDa	1	0	0	0	0
595	putative cation transport protein	630_gi|115249713	24 kDa	0	1	0	0	0
596	HfQ protein; RNA-binding protein (host factor-I protein)	630_gi|115251020	10 kDa	0	1	0	0	0
598	rod shape-determining protein	630_gi|115249136	36 kDa	1	0	0	0	0
599	putative cytidine and deoxycytidylate deaminase	630_gi|115252537	16 kDa	1	0	0	0	0
616	conserved hypothetical protein	630_gi|115252126	20 kDa	0	1	0	0	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	54 proteins m=34, r=6, h=6, R=6, s/v=1, t=1			1	2	1	2	1
	Ebeads and SE beads							
243	putative alanyl-tRNA synthetase	630_gi|115250316	98 kDa	3	4	1	1	1
53	50S ribosomal protein L18	630_gi|115249093	14 kDa	4	4	5	5	0
54	30S ribosomal protein S17	630_gi|115249086	10 kDa	5	2	2	1	0
106	putative GTP-binding protein	630_gi|115251683	50 kDa	9	11	5	3	0
111	phosphate butyryltransferase	630_gi|115249121	32 kDa	3	4	3	2	0
120	metallo beta-lactamase superfamily protein	630_gi|115250323	62 kDa	4	4	6	4	0
126	pyruvate, phosphate dikinase	630_gi|115251462	97 kDa	9	10	3	1	0
139	transcription termination factor Rho	630_gi|115252548	61 kDa	4	4	3	3	0
151	30S ribosomal protein S21	630_gi|115251500	7 kDa	3	3	4	4	0
165	putative subunit of oxidoreductase	630_gi|115249126	27 kDa	3	4	1	0	0
169	30S ribosomal protein S18	630_gi|115252726	9 kDa	1	1	1	1	0
171	glyceraldehyde-3-phosphate dehydrogenase 1	630_gi|115250813	37 kDa	3	3	3	2	0
173	putative electron transfer protein	630_gi|115252303	47 kDa	1	0	7	7	0
182	S-adenosylmethionine synthetase	630_gi|115249139	44 kDa	1	2	4	3	0
184	ptsI, 2998875-2997163 (CounterClockwise) Hypothetical Protein ptsI phosphoenolpyruvate-protein phosphotransferase	stt_BASYS02857 115251808	63 kDa	5	4	1	1	0

186	3-oxoacyl-acyl-carrier-protein synthase II	630_gi|115250217	44 kDa	5	4	3	3	0
189	50S ribosomal protein L32	630_gi|115250209	6 kDa	0	1	1	1	0
193	putative ornithine cyclodeaminase	630_gi|115249560	36 kDa	2	2	2	2	0
208	conserved hypothetical protein	630_gi|115252730	7 kDa	0	1	4	3	0
213	glucosamine--fructose-6-phosphate aminotransferase isomerizing	630_gi|115249129	67 kDa	3	5	1	1	0
214	GntR-family transcriptional regulator	630_gi|115252628	25 kDa	3	6	1	1	0
220	putative pyridine nucleotide-disulfide oxidoreductase	630_gi|115250843	59 kDa	2	3	1	1	0
225	trans-2-enoyl-ACP reductase	630_gi|115250213	33 kDa	2	3	3	1	0
234	putative molybdenum-binding subunit of oxidoreductase	630_gi|115251153	82 kDa	0	1	7	4	0
236	50S ribosomal protein L23	630_gi|115249079	11 kDa	0	1	4	4	0
257	ycjV, 3063107-3064216 (Clockwise) Hypothetical Protein ycjV ABC transporter, ATP-binding protein	sta_BASYS02921 115251510	42 kDa	1	1	2	2	0
258	putative membrane protein	630_gi|115250942	14 kDa	1	1	1	1	0
260	conserved hypothetical protein	630_gi|115250341	18 kDa	1	1	1	1	0
264	ppaC, 1051638-1053236 (Clockwise) Probable manganese-dependent inorganic pyrophosphatase manganese-dependent inorganic pyrophosphatase	stb_BASYS01013 115249342	59 kDa	3	5	1	0	0
267	chaperone protein (heat shock protein)	630_gi|115249282	75 kDa	3	5	1	0	0
277	conserved hypothetical protein	630_gi|115251019	48 kDa	1	1	1	1	0
292	anaerobic ribonucleoside-triphosphate reductase	630_gi|115249116	89 kDa	1	2	2	1	0
296	uridylate kinase	630_gi|115251192	26 kDa	1	2	1	1	0
299	putative glycosyl transferase	630_gi|115249022	45 kDa	0	1	3	2	0
320	GMP synthase glutamine-hydrolyzing	630_gi|115249206	57 kDa	3	3	0	1	0
327	putative small-molecule-binding protein	630_gi|115249015	20 kDa	1	1	1	1	0
346	RRF2-family transcriptional regulator	630_gi|115250312	16 kDa	0	1	2	2	0
347	V-type sodium ATP synthase subunit D	630_gi|115252010	26 kDa	0	1	2	2	0
348	putative iron-only hydrogenase,electron-transferring subunit	630_gi|115252465	18 kDa	1	1	1	1	0
362	UDP-glucose 4-epimerase	630_gi|115251765	37 kDa	2	2	1	2	0
364	putative oxidoreductase, NAD/FAD binding subunit	630_gi|115249186	46 kDa	2	1	0	1	0
371	DeoR-family transcriptional regulator (fatty acid and phospholipid biosynthesis regulator)	630_gi|115250210	21 kDa	1	1	1	0	0
373	SOS regulatory protein	630_gi|115250977	24 kDa	0	1	1	1	0

390	delta-aminolevulinic acid dehydratase	630_gi|115252479	36 kDa	1	1	1	0	0
392	(3R)-hydroxymyristoyl-acyl carrier protein dehydratase	630_gi|115249137	16 kDa	1	1	0	1	0
416	putative ATP-dependent RNA helicase	630_gi|115249778	61 kDa	1	0	1	1	0
418	putative DNA topoisomerase	630_gi|115251327	81 kDa	1	1	0	1	0
443	conserved hypothetical protein	630_gi|115250288	8 kDa	2	1	0	1	0
451	ATP-dependent Clp protease	630_gi|115249029	91 kDa	1	0	0	2	0
456	BASYS03265, 3471128-3469293 (CounterClockwise) Hypothetical Protein BASYS03265 cell surface protein	sta_BASYS03265 115251842	67 kDa	1	0	0	1	0
477	conserved hypothetical protein	630_gi|115252308	9 kDa	1	0	1	0	0
486	putative iron-only hydrogenase, catalytic subunit	630_gi|115252467	65 kDa	1	0	1	0	0
503	ATP-dependent nuclease subunit B	630_gi|115250061	134 kDa	1	0	1	0	0
622	ydeM, 3352800-3351373 (CounterClockwise) Hypothetical Protein ydeM Radical SAM-family protein	stt_BASYS03151 115249169	55 kDa	0	1	1	0	0
	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	14 proteins m=5, r=1, h=4, R=3, s/v=1, t=0			1	2	1	2	1
	SE beads and crude							
97	conserved hypothetical protein	630_gi|115249020	12 kDa	0	0	6	6	3
237	elongation factor P	630_gi|115250279	21 kDa	0	0	1	0	4
238	50S ribosomal protein L30	630_gi|115249095	7 kDa	0	0	1	1	1
265	BASYS00644, 660422-660628 (Clockwise) Hypothetical Protein LJ conserved hypothetical protein	sta_BASYS00644 115249289	8 kDa	0	0	3	2	1
269	putative tRNA binding protein	630_gi|115251499	17 kDa	0	0	1	0	4
306	cell surface protein (putative S-layer protein precursor)	630_gi|115251844	66 kDa	0	0	2	3	1
313	tellurium resistance protein	630_gi|115250675	21 kDa	0	0	1	1	1
314	conserved hypothetical protein	630_gi|115249751	7 kDa	0	0	1	0	3
389	transcription antitermination protein	630_gi|115249064	20 kDa	0	0	1	1	1
413	conserved hypothetical protein	630_gi|115250440	24 kDa	0	0	0	1	1
419	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	630_gi|115249132	45 kDa	0	0	1	1	1
450	penicillinase repressor	630_gi|115249484	15 kDa	0	0	0	1	2
465	putative tRNA (Uracil-5-)-methyltransferase	630_gi|115249412	51 kDa	0	0	1	0	1
475	putative histidinol-phosphate aminotransferase	630_gi|115250590	40 kDa	0	0	0	1	1

	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	58 proteins m=38, r=2, h=7, R=5, s/v=3, t=3			1	2	1	2	1
	SE beads only							
170	RNA polymerase sigma factor (sigma-43)	630_gi|115250496	44 kDa	0	0	5	6	0
215	putative FAD-binding subunit of oxidoreductase	630_gi|115251155	30 kDa	0	0	6	4	0
254	translation initiation factor IF-1	630_gi|115249101	8 kDa	0	0	2	1	0
280	putative helicase	630_gi|115250059	87 kDa	0	0	1	1	0
283	putative thiol peroxidase (bacterioferritin comigratory protein)	630_gi|115250867	18 kDa	0	0	4	4	0
284	riboflavin biosynthesis protein	630_gi|115250350	35 kDa	0	0	4	3	0
295	50S ribosomal protein L31	630_gi|115252547	8 kDa	0	0	2	2	0
301	ribosome-binding factor A	630_gi|115250346	14 kDa	0	0	2	3	0
309	UDP-N-acetylglucosamine--N-acetylmuramyl-(penta peptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	630_gi|115251702	45 kDa	0	0	3	2	0
310	coenzyme A biosynthesis bifunctional protein	630_gi|115251643	44 kDa	0	0	2	2	0
330	RecA protein (recombinase A)	630_gi|115250364	37 kDa	0	0	2	2	0
333	putative formate dehydrogenase	630_gi|115251232	83 kDa	0	0	2	1	0
338	putative nogalamycin resistance protein	630_gi|115252273	83 kDa	0	0	1	1	0
355	preprotein translocase SecA subunit	630_gi|115249152	102 kDa	0	0	1	3	0
357	putative cell wall hydrolase	630_gi|115250166	45 kDa	0	0	1	2	0
365	putative mannosyl-glycoprotein endo-beta-N-acetylglucosamidase	630_gi|115250339	66 kDa	0	0	2	2	0
366	fatty acid/phospholipid synthesis protein	630_gi|115250211	37 kDa	0	0	2	2	0
370	putative capsular polysaccharide biosynthesis glycosyl transferase	630_gi|115251823	41 kDa	0	0	3	2	0
374	putative toxic anion resistance protein	630_gi|115251391	44 kDa	0	0	1	3	0
375	PTS system, glucitol/sorbitol-specific IIb component	630_gi|115249782	13 kDa	0	0	1	2	0
394	electron transport complex protein	630_gi|115250168	48 kDa	0	0	2	1	0
395	probable GTP-binding protein	630_gi|115252356	22 kDa	0	0	2	1	0
396	putative ATP-dependent RNA helicase	630_gi|115251160	43 kDa	0	0	1	2	0
397	ATP-dependent protease La	630_gi|115252357	90 kDa	0	0	1	2	0
401	putative glycosyltransferase	630_gi|115250250	26 kDa	0	0	2	1	0
420	conserved hypothetical protein	630_gi|115252367	16 kDa	0	0	1	1	0
421	conserved hypothetical protein	630_gi|115252542	39 kDa	0	0	1	1	0
449	iron-dependent hydrogenase	630_gi|115249907	56 kDa	0	0	2	1	0
453	stage V sporulation protein S	630_gi|115250981	9 kDa	0	0	1	1	0
454	TetR-family transcriptional regulator	630_gi|115251558	35 kDa	0	0	1	1	0

455	bifunctional protein includes: UDP-N-acetylglucosamine pyrophosphorylase and glucosamine-1-phosphate N-acetyltransferase	630_gi|115252576	50 kDa	0	0	1	1	0
479	putative carbon monoxide dehydrogenase accessory protein	630_gi|115249734	29 kDa	0	0	1	0	0
491	conserved hypothetical protein	630_gi|115252624	16 kDa	0	0	1	1	0
492	putative ferrous iron transport protein A	630_gi|115250789	8 kDa	0	0	1	1	0
497	queuine tRNA-ribosyltransferase	630_gi|115251855	43 kDa	0	0	2	0	0
504	carbon storage regulator	630_gi|115249242	8 kDa	0	0	1	0	0
505	putative flagellar protein	630_gi|115249264	8 kDa	0	0	1	0	0
506	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	630_gi|115250983	32 kDa	0	0	1	0	0
507	conserved hypothetical protein	630_gi|115251395	10 kDa	0	0	1	0	0
542	Spo0B-associated GTP-binding protein	630_gi|115250196	47 kDa	0	0	0	1	0
543	fabD, 1551856-1552806 (Clockwise) Hypothetical Protein fabD malonyl CoA-acyl carrier protein transacylase	sta_BASYS01457 115250214	34 kDa	0	0	0	1	0
544	xylR, 997988-997017 (CounterClockwise) Hypothetical Protein xylR -	stb_BASYS00966 -	38 kDa	0	0	0	1	0
545	undecaprenyl pyrophosphate synthetase	630_gi|115251189	28 kDa	0	0	1	0	0
546	GTP-binding elongation factor	630_gi|115251521	67 kDa	0	0	1	0	0
547	porphyrin biosynthesis protein includes: uroporphyrinogen-III methyltransferase and uroporphyrinogen-III synthase	630_gi|115252480	56 kDa	0	0	1	0	0
548	putative RNA-binding protein	630_gi|115252556	10 kDa	0	0	1	0	0
549	dimethyladenosine transferase	630_gi|115252584	32 kDa	0	0	1	0	0
567	putative radical SAM superfamily protein	630_gi|115251502	50 kDa	0	0	0	1	0
585	putative methylenetetrahydrofolate reductase	630_gi|115249739	32 kDa	0	0	0	1	0
586	putative GTP pyrophosphokinase	630_gi|115251796	85 kDa	0	0	1	0	0
588	conserved hypothetical protein	630_gi|115249056	15 kDa	0	0	1	0	0
597	two-component sensor histidine kinase	630_gi|115250874	102 kDa	0	0	1	0	0
602	excinuclease ABC subunit A	630_gi|115252471	105 kDa	0	0	1	0	0
608	putative iron ABC transporter, ATP-binding protein	630_gi|115252591	37 kDa	0	0	0	1	0
615	conserved hypothetical protein	630_gi|115249635	9 kDa	0	0	1	0	0
617	ABC transporter, ATP-binding protein	630_gi|115250573	28 kDa	0	0	0	1	0
618	putative antibiotic resistance ABC transporter,ATP-binding protein	630_gi|115251650	63 kDa	0	0	0	1	0
620	L-lysine 2,3-aminomutase	630_gi|115251307	49 kDa	0	0	0	1	0

	Identification	Number	Mass	No. unique peptides				
				E beads		SE beads		Crude
	148 proteins m=89, r=0, h=36, R=10, s/v=8, t=5			1	2	1	2	1
	Crude only							
131	toxin A	630_gi|115249677	308 kDa	0	0	0	0	20
156	probable carbohydrate hydrolase (N-terminus)	630_gi|115252061	11 kDa	0	0	0	0	6
159	10 kDa chaperonin	630_gi|115249203	10 kDa	0	0	0	0	2
166	hisC, 2820848-2821951 (Clockwise) Hypothetical Protein hisC putative histidinol-phosphate aminotransferase	stb_BASYS02725 115251555	42 kDa	0	0	0	0	16
190	putative imidazole glycerol phosphate synthase subunit	630_gi|115250592	23 kDa	0	0	0	0	4
196	ycjL, 3792708-3791962 (CounterClockwise) Hypothetical Protein ycjL putative glutamine amidotransferase	sta_BASYS03542 115252147	28 kDa	0	0	0	0	3
209	thioredoxin	630_gi|115251408	12 kDa	0	0	0	0	3
228	putative imidazole glycerol phosphate synthase subunit	630_gi|115250594	28 kDa	0	0	0	0	6
239	conserved hypothetical protein	630_gi|115251115	24 kDa	0	0	0	0	7
240	peptidyl-prolyl cis-trans isomerase	630_gi|115249340	19 kDa	0	0	0	0	6
249	transcription elongation factor	630_gi|115252616	18 kDa	0	0	0	0	4
255	inosine-5'-monophosphate dehydrogenase	630_gi|115251390	55 kDa	0	0	0	0	5
256	imidazoleglycerol-phosphate dehydratase	630_gi|115250591	22 kDa	0	0	0	0	2
262	putative formyltransferase	630_gi|115249858	44 kDa	0	0	0	0	6
263	conserved hypothetical protein	630_gi|115252018	11 kDa	0	0	0	0	4
271	probable amino-acid ABC transporter,substrate-binding protein	630_gi|115251227	29 kDa	0	0	0	0	7
272	putative reductase	630_gi|115249854	20 kDa	0	0	0	0	5
273	probable amino-acid ABC transporter,substrate-binding protein	630_gi|115251230	30 kDa	0	0	0	0	5
279	putative flavodoxin	630_gi|115249827	18 kDa	0	0	0	0	3
287	alanine racemase	630_gi|115252523	43 kDa	0	0	0	0	5
288	putative 1-(5-phosphoribosyl)-5-(5-phosphoribosylamino)methylidene amino imidazole-4-carboxamide isomerase	630_gi|115250593	27 kDa	0	0	0	0	5
303	putative lipoprotein	630_gi|115249561	23 kDa	0	0	0	0	5
304	conserved hypothetical protein	630_gi|115250755	12 kDa	0	0	0	0	3
315	putative carbon-nitrogen hydrolase	630_gi|115249501	31 kDa	0	0	0	0	6
316	histidine biosynthesis bifunctional protein includes: phosphoribosyl-AMP cyclohydrolase and phosphoribosyl-ATP pyrophosphatase	630_gi|115250595	26 kDa	0	0	0	0	5

317	putative aliphatic sulfonates ABC transporter,substrate-binding lipoprotein	630_gi|115250525	36 kDa	0	0	0	0	4
324	DNA-directed RNA polymerase omega chain	630_gi|115251644	10 kDa	0	0	0	0	4
325	conserved hypothetical protein	630_gi|115251027	11 kDa	0	0	0	0	2
341	cell surface protein (putative hemagglutinin/adhesin)	630_gi|115249532	167 kDa	0	0	0	0	4
342	uracil-DNA glycosylase	630_gi|115251535	26 kDa	0	0	0	0	4
343	conserved hypothetical protein	630_gi|115251313	13 kDa	0	0	0	0	4
344	putative FolD bifunctional protein includes: methylenetetrahydrofolate dehydrogenase; methenyltetrahydrofolate cyclohydrolase	630_gi|115249737	31 kDa	0	0	0	0	4
345	tellurium resistance protein	630_gi|115250845	21 kDa	0	0	0	0	3
359	putative peptidyl-prolyl isomerase	630_gi|115250393	28 kDa	0	0	0	0	5
360	putative regulatory protein	630_gi|115250940	19 kDa	0	0	0	0	4
361	putative permease	630_gi|115251749	17 kDa	0	0	0	0	2
380	yafV, 3404477-3403638 (CounterClockwise) Hypothetical Protein yafV putative carbon-nitrogen hydrolase	sta_BASYS03214 115251789	32 kDa	0	0	0	0	4
381	BASYS02855, 2995211-2994705 (CounterClockwise) Hypothetical Protein BF conserved hypothetical protein	sta_BASYS02855 115251448	20 kDa	0	0	0	0	4
382	putative aromatic amino acid aminotransferase	630_gi|115251881	45 kDa	0	0	0	0	3
383	anti-sigma-B factor antagonist	630_gi|115249012	12 kDa	0	0	0	0	3
384	conserved hypothetical protein	630_gi|115250810	27 kDa	0	0	0	0	2
385	conserved hypothetical protein	630_gi|115249410	13 kDa	0	0	0	0	2
386	PTS system, IIb component	630_gi|115252122	18 kDa	0	0	0	0	2
387	conserved hypothetical protein	630_gi|115249867	8 kDa	0	0	0	0	1
393	DNA mismatch repair protein	630_gi|115251023	109 kDa	1	1	1	1	0
404	BASYS01005, 1036990-1042047 (Clockwise) Hypothetical Protein BF conserved hypothetical protein	stb_BASYS01005 115249349	195 kDa	0	0	0	0	4
405	putative ATP-binding protein	630_gi|115250457	25 kDa	0	0	0	0	3
406	putative acetyltransferase	630_gi|115250223	20 kDa	0	0	0	0	3
407	V-type sodium ATP synthase subunit B	630_gi|115252011	51 kDa	0	0	0	0	3
408	conserved hypothetical protein	630_gi|115251601	68 kDa	0	0	0	0	3
409	glucose-6-phosphate isomerase	630_gi|115252341	51 kDa	0	0	0	0	2
410	fructose-6-phosphate aldolase 2	630_gi|115251164	25 kDa	0	0	0	0	2
411	PTS system, IIc component	630_gi|115252070	39 kDa	0	0	0	0	2
412	histidine triad nucleotide-binding protein	630_gi|115251501	13 kDa	0	0	0	0	2
414	hypothetical protein	630_gi|115249753	14 kDa	0	0	0	0	1

430	BASYS02132, 2224437-2224243 (CounterClockwise) Hypothetical Protein BASYS02132 -	sta_BASYS02132 -	8 kDa	0	0	0	0	3
431	aspartate carbamoyltransferase catalytic chain	630_gi|115249194	35 kDa	0	0	0	0	2
432	PTS system, IIb component	630_gi|115249216	11 kDa	0	0	0	0	2
433	putative phosphoribosylaminoimidazole-succinocarboxamide synthase	630_gi|115249503	25 kDa	0	0	0	0	2
434	PTS system, galactitol-specific IIb component	630_gi|115251381	10 kDa	0	0	0	0	2
435	hypothetical protein	630_gi|115251805	26 kDa	0	0	0	0	2
436	putative membrane protein	630_gi|115252518	39 kDa	0	0	0	0	2
437	folB, 1546460-1546855 (Clockwise) Hypothetical Protein folB dihydroneopterin aldolase	stt_BASYS01511 115250492	15 kDa	0	0	0	0	2
438	putative ATP phosphoribosyltransferase	630_gi|115250589	23 kDa	0	0	0	0	2
439	4-aminobutyrate aminotransferase	630_gi|115251211	48 kDa	0	0	0	0	2
440	cysteine desulfurase	630_gi|115250313	44 kDa	0	0	0	0	2
441	putative glyoxalase	630_gi|115252675	15 kDa	0	0	0	0	1
444	conserved hypothetical protein	630_gi|115251130	42 kDa	0	0	0	0	1
445	putative transferase	630_gi|115251800	18 kDa	0	0	0	0	1
446	conserved hypothetical protein	630_gi|115252732	22 kDa	0	0	0	0	1
447	BASYS00311, 304950-305525 (Clockwise) Hypothetical Protein BASYS00311 putative exported protein	sta_BASYS00311 115252734	21 kDa	0	0	0	0	1
466	dihydroorotate dehydrogenase electron transfer subunit	630_gi|115249195	26 kDa	0	0	0	0	2
467	putative phosphatidylethanolamine-binding regulatory protein	630_gi|115252223	18 kDa	0	0	0	0	2
468	hypothetical protein	630_gi|115250585	8 kDa	0	0	0	0	2
469	conserved hypothetical protein	630_gi|115249879	33 kDa	0	0	0	0	2
470	putative membrane protein	630_gi|115249847	43 kDa	0	0	0	0	2
471	orotidine 5'-phosphate decarboxylase	630_gi|115252656	26 kDa	0	0	0	0	2
472	conserved hypothetical protein	630_gi|115252296	17 kDa	0	0	0	0	1
473	putative glutamyl-aminopeptidase	630_gi|115251205	39 kDa	0	0	0	0	2
481	putative phage regulatory protein	630_gi|115251071	14 kDa	0	0	0	0	1
482	peptide deformylase 2	630_gi|115251641	16 kDa	0	0	0	0	1
483	putative nitroreductase	630_gi|115252261	22 kDa	0	0	0	0	1
484	pilin	630_gi|115252574	18 kDa	0	0	0	0	1
493	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	630_gi|115249051	17 kDa	0	0	0	0	2
494	putative GTP cyclohydrolase I	630_gi|115250490	21 kDa	0	0	0	0	2
498	formylglycinamidine ribonucleotide synthetase	630_gi|115249233	141 kDa	0	0	0	0	2
501	aspartokinase	630_gi|115251108	43 kDa	0	0	0	0	2
508	glutamyl-tRNA synthetase	630_gi|115249054	57 kDa	0	0	0	0	1
509	PTS system, IId component	630_gi|115249298	31 kDa	0	0	0	0	1

510	putative penicillin-binding protein repressor	630_gi|115249563	14 kDa	0	0	0	0	1
511	MarR-family transcriptional regulator	630_gi|115249852	17 kDa	0	0	0	0	1
512	conserved hypothetical protein	630_gi|115249911	14 kDa	0	0	0	0	1
513	putative flavoprotein	630_gi|115250499	21 kDa	0	0	0	0	1
514	putative molybdenum cofactor biosynthesis protein	630_gi|115250756	18 kDa	0	0	0	0	1
515	conserved hypothetical protein	630_gi|115250817	14 kDa	0	0	0	0	1
516	conserved hypothetical protein	630_gi|115250980	14 kDa	0	0	0	0	1
517	conserved hypothetical protein	630_gi|115251427	11 kDa	0	0	0	0	1
518	deoxyuridine 5'-triphosphate nucleotidohydrolase	630_gi|115251455	16 kDa	0	0	0	0	1
519	guanylate kinase	630_gi|115251645	23 kDa	0	0	0	0	1
520	putative protein translocase subunit	630_gi|115251854	11 kDa	0	0	0	0	1
521	putative flavodoxin	630_gi|115251878	20 kDa	0	0	0	0	1
522	V-type sodium ATP synthase subunit K	630_gi|115252016	16 kDa	0	0	0	0	1
523	putative PTS syste, IIa component	630_gi|115252157	18 kDa	0	0	0	0	1
524	hypoxanthine phosphoribosyltransferase	630_gi|115252288	20 kDa	0	0	0	0	1
525	conserved hypothetical protein	630_gi|115252297	29 kDa	0	0	0	0	1
526	ATP synthase epsilon chain (partial)	630_gi|115252527	9 kDa	0	0	0	0	1
527	ATP synthase B chain	630_gi|115252532	20 kDa	0	0	0	0	1
528	BASYS02157, 2250534-2250301 (CounterClockwise) Hypothetical Protein BASYS02157 -	sta_BASYS02157 -	9 kDa	0	0	0	0	1
529	putative phosphoglycerate mutase	630_gi|115250859	24 kDa	0	0	0	0	1
530	putative exported protein	630_gi|115249010	11 kDa	0	0	0	0	1
531	6-phospho-beta-glucosidase	630_gi|115252149	55 kDa	0	0	0	0	1
533	V-type sodium ATP synthase subunit E	630_gi|115252015	21 kDa	0	0	0	0	1
534	putative membrane protein	630_gi|115251750	24 kDa	0	0	0	0	1
536	BASYS02919, 3062106-3060631 (CounterClockwise) Hypothetical Protein BASYS02919 hypothetical protein	sta_BASYS02919 115251508	58 kDa	0	0	0	0	1
538	conserved hypothetical protein	630_gi|115251465	46 kDa	0	0	0	0	1
540	PTS system, IIa component	630_gi|115251622	17 kDa	0	0	0	0	1
550	4-hydroxyphenylacetate decarboxylase, regulatory subunit	630_gi|115249163	10 kDa	0	0	0	0	1
551	bifunctional purine biosynthesis protein includes: phosphoribosylaminoimidazolecarboxamide formyltransferase and IMP cyclohydrolase	630_gi|115249231	57 kDa	0	0	0	0	1
552	putative PTS system, IIa component	630_gi|115249295	15 kDa	0	0	0	0	1
553	hypothetical protein	630_gi|115249604	92 kDa	0	0	0	0	1
554	arginyl-tRNA synthetase	630_gi|115249727	65 kDa	0	0	0	0	1
555	putative hydrolase	630_gi|115249857	33 kDa	0	0	0	0	1

556	branched chain amino acid transport system carrier protein	630_gi|115250294	45 kDa	0	0	0	0	1
557	putative phage protein	630_gi|115250400	39 kDa	0	0	0	0	1
558	putative phosphoglucomutase	630_gi|115250617	64 kDa	0	0	0	0	1
559	putative protein export chaperone	630_gi|115251317	43 kDa	0	0	0	0	1
560	conserved hypothetical protein	630_gi|115251507	39 kDa	0	0	0	0	1
561	low-specificity L-threonine aldolase	630_gi|115251648	38 kDa	0	0	0	0	1
562	putative amidohydrolas	630_gi|115251756	51 kDa	0	0	0	0	1
563	conserved hypothetical protein	630_gi|115252306	21 kDa	0	0	0	0	1
564	putative nitroreductase	630_gi|115252412	20 kDa	0	0	0	0	1
565	BASYS02134, 2227250-2227441 (Clockwise) Hypothetical Protein BASYS02134 -	sta_BASYS02134 -	8 kDa	0	0	0	0	1
566	BASYS03517, 3762887-3762459 (CounterClockwise) Hypothetical Protein BASYS03517 -	sta_BASYS03517 -	17 kDa	0	0	0	0	1
581	probable carbohydrate hydrolase	630_gi|115252060	42 kDa	0	0	0	0	1
582	putative carbonic anhydrase	630_gi|115251267	24 kDa	0	0	0	0	1
584	purine nucleoside phosphorylase	630_gi|115250257	29 kDa	0	0	0	0	1
587	putative molybdenum ABC transporter,substrate-binding protein	630_gi|115249883	29 kDa	0	0	0	0	1
590	putative acetyltransferase	630_gi|115251215	20 kDa	0	0	0	0	1
591	dihydrodipicolinate reductase	630_gi|115252286	28 kDa	0	0	0	0	1
592	putative membrane protein	630_gi|115249011	10 kDa	0	0	0	0	1
593	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	630_gi|115252228	56 kDa	0	0	0	0	1
600	conserved hypothetical protein	630_gi|115252731	10 kDa	0	0	0	0	1
601	cell surface protein	630_gi|115249861	34 kDa	0	0	0	0	1
603	electron transport complex protein	630_gi|115250172	20 kDa	0	0	0	0	1
605	V-type sodium ATP synthase subunit G	630_gi|115252013	12 kDa	0	0	0	0	1
612	signal peptidase I	630_gi|115249570	20 kDa	0	0	0	0	1
613	amino acid ABC transporter, substrate-binding protein	630_gi|115249767	30 kDa	0	0	0	0	1
621	putative amidohydrolase	630_gi|115250653	43 kDa	0	0	0	0	1